

Durham Research Online

Deposited in DRO:

07 July 2020

Version of attached file:

Published Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Pochiraju, Bhimasankaram and Seshadri, Sridhar and Thomakos, Dimitrios D. and Nikolopoulos, Konstantinos (2020) 'Non-negativity of a quadratic form with applications to panel data estimation, forecasting and optimization.', *Stats.*, 3 (3). pp. 185-202.

Further information on publisher's website:

<https://doi.org/10.3390/stats3030015>

Publisher's copyright statement:

Additional information:

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

Article

Non-Negativity of a Quadratic form with Applications to Panel Data Estimation, Forecasting and Optimization

Bhimasankaram Pochiraju ^{1,†}, Sridhar Seshadri ² and Dimitrios D. Thomakos ³
and Konstantinos Nikolopoulos ^{4,*}

¹ Indian School of Business, Gachibowli, Hyderabad, Telangana 500032, India

² Gies College of Business, University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA; sridhar@illinois.edu

³ Department of Economics, School of Management and Economics, University of Peloponnese, 22100 Peloponnese, Greece; thomakos@uop.gr

⁴ Durham University Business School, Durham DH1 3LB, UK

* Correspondence: kostas.nikolopoulos@durham.ac.uk or kostas@fortank.com

† Bhimasankaram Pochiraju was Professor at the Indian School of Business in India. He passed away on 1 April 2018. He and Sridhar Seshadri worked on this paper when they were at the Indian School of Business.

Received: 26 May 2020; Accepted: 30 June 2020; Published: 6 July 2020



Abstract: For a symmetric matrix \mathbf{B} , we determine the class of \mathbf{Q} such that $\mathbf{Q}^t\mathbf{B}\mathbf{Q}$ is non-negative definite and apply it to panel data estimation and forecasting: the Hausman test for testing the endogeneity of the random effects in panel data models. We show that the test can be performed if the estimated error variances in the fixed and random effects models satisfy a specific inequality. If it fails, we discuss the restrictions under which the test can be performed. We show that estimators satisfying the inequality exist. Furthermore, we discuss an application to a constrained quadratic minimization problem with an indefinite objective function.

Keywords: quadratic form; Non-negativity; Hausman test; optimization

JEL Classification: C01; C02; C61

1. Introduction

Optimization of quadratic structures and corrections to the construction of covariance matrices has a long history in econometrics and financial economics. From the early stages of simultaneous equations modeling, to testing panel data models, to recent advances in handling volatilities and correlations of financial returns in large systems, to applications in portfolio management, all these issues have contained a component of either quadratic structures and covariance estimation or some form of optimization based on them. The importance of the above in forecasting cannot be overstated as they all relate to decision-making at some future time period: a good panel data model can be used in generating out-of-sample forecasts, a well-constructed covariance matrix can be used in optimizing the weights of a portfolio or for building a model for volatility and correlation forecasting. Although we rarely admit it, most models and mathematical operations are –in the end –geared for forecasting.

In this paper we present a novel mathematical approach for solving a particular class of quadratic optimization problems with applications in econometrics, statistics and portfolio construction. At the center of these mathematical derivations is a symmetric indefinite matrix, in applications a covariance matrix. The indefinite nature of this matrix can come from many sources but the common

ground to all is rank deficiency or rank indeterminacy based on redundant information in the variables from which we compute that said matrix: a large number of variables involved or some deficiency in the structure of the underlying problem. The problem of this indeterminacy leads to other problems in different set-ups: in the context of a matrix-version of the well-known Hausman test in econometrics a difference in covariance matrices required for the application of the test might not be positive definite; in the context of a large portfolio optimization the covariance matrix of the financial returns might not be positive definite. Admittedly, there are many solutions that have been proposed for these problems; for example, the Hausman test can be computed with a regression approach so that no matrix inversion is required (or a generalized inverse can be used, as has been done in many parts of the relevant literature); in the context of a rank-deficient covariance there are a number of solutions that have been proposed that correct this. So, what are the new insights that this paper has to offer? First, in the context of the Hausman test the results offer a very clear rule as to when the test (in its matrix form) can be applied. This is useful as a specification tool, and is very simple to compute and report—so when the underlying conditions do not hold for the test to be performed in its matrix form some caution is warranted, both in model specification and model performance. Second, in the context of portfolio optimization the results suggest a particular procedure for optimizing a portfolio given a set of particular constraints without reference to the potential problems of rank indeterminacy of the covariance matrix. This therefore allows for a direct solution to the optimization problem even with sample covariance matrices, computed with a limited number of observations. Finally, it should be noted that the optimization problems discussed in this paper have potential for other applications as well. For example, one might consider the case of a least generalized least squares-like problem for which the covariance matrix has rank indeterminacy because of, say, less observations than variables. In that case, given constraints, the solution proposed here can be utilized. Of course such a model can then be used in forecasting.

2. Notations

In this paper, we obtain the class of all \mathbf{Q} such that $\mathbf{Q}^t \mathbf{B} \mathbf{Q}$ is non-negative definite (nnd) where \mathbf{B} is a given symmetric indefinite matrix (the terms ‘non-negative’ and ‘nonnegative’ are used interchangeably in the literature). Why is this important? Primarily because there is a very active interest in the topic in the literature recently [1–3], but even more importantly due to the vast array of applications in statistics, finance and economics, most notably in panel data econometrics and quadratic optimization, as elaborated in the subsequent sections.

(a) Let θ be a parametric vector of interest and \mathbf{t} be an unbiased estimator of θ , the dispersion matrix of which depends on some parameters γ . The estimated dispersion matrix $\hat{D}(\mathbf{t})$ based on $\hat{\gamma}$ may turn out to be an indefinite matrix. It is of interest to find out the class of all linear parametric functions $\mathbf{Q}^t \theta$ for which the estimated dispersion matrix of the unbiased estimator $\mathbf{Q}^t \mathbf{t}$ is non-negative definite. Take a specific instance where $\hat{D}(\mathbf{t}) = \sigma^2((1 - \hat{\rho})\mathbf{I} + \hat{\rho}\mathbf{1}\mathbf{1}^t)$ and let $\hat{\rho} < -\frac{1}{n-1}$ where \mathbf{t} is of order $n \times 1$. In this case $\hat{D}(\mathbf{t})$ is indefinite.

(b) Again, let θ be a parametric vector of interest and let \mathbf{t}_1 and \mathbf{t}_2 be two unbiased estimators of θ . We say that \mathbf{t}_1 is superior to \mathbf{t}_2 if $\hat{D}(\mathbf{t}_2) - \hat{D}(\mathbf{t}_1)$ is non-negative definite or equivalently if $\hat{D}(\mathbf{t}_1)$ is below $\hat{D}(\mathbf{t}_2)$ under the Löwner order [4]. Suppose neither of \mathbf{t}_1 and \mathbf{t}_2 is superior to the other. It is of interest to find out sets of linear functions $\mathbf{Q}^t \theta$ of θ such that $\hat{D}(\mathbf{t}_2) - \hat{D}(\mathbf{t}_1)$ is non-negative definite. For a specific case, we consider the fixed effects and random effects panel data models [5] to examine when the issue of endogeneity in the random effects model can be checked using the Hausman test. We study the following:

- (i) Suppose we choose and fix the functional form of the estimators of the variance components. We obtain the class of all regressor matrices for which we can perform the Hausman test.
- (ii) Suppose for given data on regressors the difference in the estimated dispersion matrices is indefinite. We obtain the class of linear compounds of the regression coefficient vector for which we can perform the Hausman test.

- (iii) We note that there always exists an estimator of the variance of the white noise part of the error in the random effects model for which the difference in the estimated dispersion matrices of the fixed effects and random effects estimators of the regression coefficient vector is indeed non-negative definite. Reference [6]’s estimator is one such estimator of the variance component mentioned above. There can be others.
- (iv) We extend the above results to the cases where either the random error or the random effects or both are heteroscedastic.

We are aware that there are alternatives to the traditional Hausman test such as in Reference [7] that incorporates the time-invariant parts of the regressors in the random effects model. However, the most popular test to date is the traditional Hausman test.

(c) Consider the problem of minimizing $\mathbf{x}^t \mathbf{B} \mathbf{x}$ subject to $\mathbf{A} \mathbf{x} = \mathbf{0}$. Clearly, this is equivalent to the unconstrained minimization problem $\mathbf{u}^t (\mathbf{I} - \mathbf{A}^+ \mathbf{A}) \mathbf{B} (\mathbf{I} - \mathbf{A}^+ \mathbf{A}) \mathbf{u}$, which has a finite solution if and only if $(\mathbf{I} - \mathbf{A}^+ \mathbf{A}) \mathbf{B} (\mathbf{I} - \mathbf{A}^+ \mathbf{A})$ is non-negative definite. (\mathbf{A}^+ denotes the Moore-Penrose inverse of \mathbf{A} .) Thus, it is often of interest to explicitly obtain the class of all vectors \mathbf{x} such that $\mathbf{x}^t \mathbf{B} \mathbf{x} \geq 0$, where \mathbf{B} is an indefinite real symmetric matrix. Unfortunately, this class is not a subspace of \mathbb{R}^n . It is also not a convex set. In this paper, we characterize the class of all matrices \mathbf{Q} such that $\mathbf{Q}^t \mathbf{B} \mathbf{Q}$ is non-negative definite (nnd). We then study the problem of minimization of a quadratic form $\mathbf{x}^t \mathbf{B} \mathbf{x}$ subject to $\mathbf{A} \mathbf{x} = \mathbf{b}$, where \mathbf{B} is an indefinite matrix and \mathbf{b} is in the column space of \mathbf{A} . Given a matrix \mathbf{A} , we characterize the class of all real symmetric matrices \mathbf{B} and vectors \mathbf{b} in the column space of \mathbf{A} for which the aforementioned problem has a finite solution. It turns out that one of the key conditions for the minimization problem to have a finite solution is the non-negative definiteness of $\mathbf{Q}^t \mathbf{B} \mathbf{Q}$ for a suitable orthogonal projection matrix \mathbf{Q} .

In Section 3, we state the results on non-negative definite matrices and generalized inverses which will be needed in the later sections. In Section 4, we obtain necessary and sufficient conditions for $\mathbf{Q}^t \mathbf{B} \mathbf{Q}$ to be non-negative definite (when \mathbf{B} is a symmetric indefinite matrix). Based on this, we develop an algorithm to generate all such matrices \mathbf{Q} . We then specialize to the cases where \mathbf{A} has (i) just one negative eigenvalue and (ii) just one positive eigenvalue. As a special case of (i), we consider the intraclass correlation matrix which comes naturally as the dispersion matrix in random effects models. In Section 6, we study, in detail, the issues related to performing Hausman test mentioned in (b) above. In Section 7, we show that the problem of finding the class of all matrices \mathbf{Q} such that $\mathbf{Q}^t \mathbf{B} \mathbf{Q}$ is nnd is equivalent to the solution of the quadratic optimization problem: Minimize $\mathbf{x}^t \mathbf{B} \mathbf{x}$ subject to $\mathbf{A} \mathbf{x} = \mathbf{0}$ varying over matrices \mathbf{A} . As we shall show, the connection between \mathbf{Q} and \mathbf{A} comes through the relationship, null space of \mathbf{A} being equal to the column space of \mathbf{Q} . Given \mathbf{A} , we then determine the class of all matrices \mathbf{Q} such that $\mathcal{N}(\mathbf{A}) = \mathcal{C}(\mathbf{Q})$. Likewise, given \mathbf{Q} , we also determine the class of all matrices \mathbf{A} such that $\mathcal{N}(\mathbf{A}) = \mathcal{C}(\mathbf{Q})$. In Section 8, we study, in some detail, the constrained optimization problem of minimizing $\mathbf{x}^t \mathbf{B} \mathbf{x}$ subject to $\mathbf{A} \mathbf{x} = \mathbf{b}$ where $\mathbf{b} \in \mathcal{C}(\mathbf{A})$ and \mathbf{B} is a symmetric indefinite matrix. We consider two cases: (i) the problem has a solution for some non-null vector $\mathbf{b} \in \mathcal{C}(\mathbf{A})$ and (ii) the problem has a solution for every non-null $\mathbf{b} \in \mathcal{C}(\mathbf{A})$. Finally, Section 9 concludes.

We use real vectors and matrices in this paper and use the following notations. For a matrix \mathbf{A} , $\rho(\mathbf{A})$, $\text{tr}(\mathbf{A})$, $\mathcal{C}(\mathbf{A})$, $\mathcal{N}(\mathbf{A})$, \mathbf{A}^t , \mathbf{A}^- , \mathbf{A}^+ , $P_{\mathbf{A}}$ denote respectively the rank, trace, column space, null space, transpose, generalized inverse, Moore-Penrose inverse and orthogonal projector into the column space of a matrix \mathbf{A} . For a positive integer r , $\mathbf{1}_r$ denotes a column vector with r components where each component is 1. Further, $\bar{\mathbf{J}}_r$ denotes the matrix of order $r \times r$ each element of which is $\frac{1}{r}$. Clearly $\bar{\mathbf{J}}_r = P_{\mathbf{1}_r}$. The orthogonal projector $\mathbf{I}_r - \bar{\mathbf{J}}_r$ is denoted by \mathbf{E}_r . For matrices \mathbf{A} and \mathbf{B} , $\mathbf{A} \otimes \mathbf{B}$ denotes the Kronecker product defined as $((a_{ij} \mathbf{B}))$. A symmetric matrix \mathbf{A} is said to be non-negative definite (nnd) if $\mathbf{x}^t \mathbf{A} \mathbf{x} \geq 0$

for all vectors \mathbf{x} . The symbol $\text{diag}(\mathbf{A}, \mathbf{B}, \mathbf{C})$ denotes a block diagonal matrix $\begin{pmatrix} \mathbf{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{C} \end{pmatrix}$. For a random vector ϵ , $E(\epsilon)$ and $D(\epsilon)$ denote the expectation vector and the dispersion matrix of ϵ . Also $\text{cov}(\alpha, \xi)$ denotes the covariance matrix of α with ξ .

3. Preliminaries

In this section, we provide a few results which are well-known and which will be used in the later sections of this paper.

Lemma 1. Let $\mathbf{M} = \begin{pmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{Q}^t & \mathbf{S} \end{pmatrix}$ be a real symmetric matrix where \mathbf{P} and \mathbf{S} are real symmetric matrices. Then \mathbf{M} is nnd if and only if

- (i) \mathbf{P} is nnd,
- (ii) $\mathcal{C}(\mathbf{Q}) \subseteq \mathcal{C}(\mathbf{P})$, and
- (iii) $\mathbf{S} - \mathbf{Q}^t \mathbf{P}^- \mathbf{Q}$ is nnd.

(In view of (ii), $\mathbf{Q}^t \mathbf{P}^- \mathbf{Q}$ is invariant under choices of generalized inverses of \mathbf{P} .)

The following lemma is well known. For a proof, please see Reference [8].

Lemma 2. Let \mathbf{A} and \mathbf{B} be matrices of order $m \times n$. Then $\mathbf{A}\mathbf{A}^t = \mathbf{B}\mathbf{B}^t$ if and only if $\mathbf{A} = \mathbf{B}\mathbf{T}$, where \mathbf{T} is an orthogonal matrix.

For the proofs of the following Lemmas 3–5, please see Rao and Mitra, 1971.

Lemma 3. Let \mathbf{A} be a matrix of order $m \times n$. Let $\mathbf{b} \in \mathcal{C}(\mathbf{A})$. Let \mathbf{G} be a generalized inverse of \mathbf{A} . Then

- (i) $\mathcal{N}(\mathbf{A}) = \mathcal{C}(\mathbf{I} - \mathbf{GA})$.
- (ii) The class of all solutions to $\mathbf{Ax} = \mathbf{b}$ is given by $\mathbf{Gb} + (\mathbf{I} - \mathbf{GA})\boldsymbol{\zeta}$, where $\boldsymbol{\zeta}$ is arbitrary.

Lemma 4. Let \mathbf{A} be an $m \times n$ matrix of rank $r (> 0)$. Let $\mathbf{A} = \mathbf{U} \begin{pmatrix} \boldsymbol{\Delta} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V}^t$ be a singular value decomposition of \mathbf{A} where \mathbf{U} and \mathbf{V} are orthogonal matrices and $\boldsymbol{\Delta}$ is a positive definite (pd) diagonal matrix of order $r \times r$. Then the class of the generalized inverses of \mathbf{A} is given by $\mathbf{V} \begin{pmatrix} \boldsymbol{\Delta}^{-1} & \mathbf{L} \\ \mathbf{M} & \mathbf{N} \end{pmatrix} \mathbf{U}^t$ where $\mathbf{L}, \mathbf{M}, \mathbf{N}$ are arbitrary. In particular, the Moore-Penrose inverse \mathbf{A}^+ of \mathbf{A} is given by $\mathbf{V} \begin{pmatrix} \boldsymbol{\Delta}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{U}^t$.

Lemma 5. Let \mathbf{C} and \mathbf{D} be nnd matrices of the same order. Then there exists a nonsingular matrix \mathbf{T} such that $\mathbf{T}^t \mathbf{C} \mathbf{T}$ and $\mathbf{T}^t \mathbf{D} \mathbf{T}$ are diagonal matrices.

The following lemma on quadratic optimization is well-known.

Lemma 6. Let \mathbf{C} be a real-symmetric matrix of order $n \times n$, then the function $f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^t \mathbf{C} \mathbf{x} - \mathbf{d}^t \mathbf{x}$ has a minimum value if and only if \mathbf{C} is nnd and $\mathbf{d} \in \mathcal{C}(\mathbf{C})$, in which case the minimum value is given by $-\frac{1}{2} \mathbf{d}^t \mathbf{C}^+ \mathbf{d}$. Furthermore, if $\mathbf{C} = \mathbf{S}^t \begin{pmatrix} \boldsymbol{\Lambda} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{S}$ is a spectral decomposition of \mathbf{C} , where \mathbf{S} is orthogonal and $\boldsymbol{\Lambda}$ is diagonal positive definite matrix of order $r \times r$, then the optimal value is achieved by all vectors \mathbf{x} of the form,

$$\mathbf{x} = \mathbf{C}^+ \mathbf{d} + \mathbf{S}^t \begin{pmatrix} \mathbf{0} \\ \mathbf{z} \end{pmatrix}$$

for any $\mathbf{z} \in \mathbb{R}^{n-r}$.

The following result is well-known for researchers in parallel sums of matrices and shorted operators. For a proof, see Reference [4].

Lemma 7. Let \mathbf{A} and \mathbf{B} be nnd matrices of the same order. Then $\mathcal{C}(\mathbf{A}) \cap \mathcal{C}(\mathbf{B}) = \mathcal{C}(\mathbf{A}(\mathbf{A} + \mathbf{B})^- \mathbf{B})$, where $(\mathbf{A} + \mathbf{B})^-$ is any generalized inverse of $\mathbf{A} + \mathbf{B}$.

The following lemmas are well known. For a proof, please see Reference Rao and Bhimasankaram [8].

Lemma 8. Let \mathbf{A} and \mathbf{B} be positive definite matrices of the same order. Then $\mathbf{B}^{-1} - \mathbf{A}^{-1}$ is nnd if and only if $\mathbf{A} - \mathbf{B}$ is nnd.

Lemma 9. Let \mathbf{A} be a nonsingular matrix of order $n \times n$ and let \mathbf{u} and \mathbf{v} be vectors of order $n \times 1$. Then $\mathbf{A} + \mathbf{u}\mathbf{v}^t$ is nonsingular if and only if $1 + \mathbf{v}^t \mathbf{A}^{-1} \mathbf{u}$ is not equal to 0. Also if $1 + \mathbf{v}^t \mathbf{A}^{-1} \mathbf{u}$ is not equal to 0, then $(\mathbf{A} + \mathbf{u}\mathbf{v}^t)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{u} \mathbf{v}^t \mathbf{A}^{-1} / (1 + \mathbf{v}^t \mathbf{A}^{-1} \mathbf{u})$.

Lemma 10. Let \mathbf{A} and \mathbf{D} be nonsingular matrices of orders $n \times n$ and $r \times r$ respectively and let \mathbf{B} and \mathbf{C} be matrices of orders $n \times r$ and $r \times n$ respectively. Then $\mathbf{A} + \mathbf{BDC}$ is nonsingular if and only if $\mathbf{W} = \mathbf{D}^{-1} + \mathbf{CA}^{-1}\mathbf{B}$ is nonsingular. Also if \mathbf{W} is nonsingular, then $(\mathbf{A} + \mathbf{BDC})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{D}^{-1} + \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1}$.

Lemma 11. Let \mathbf{A} and \mathbf{B} be matrices of orders $m \times n$ and $n \times m$ respectively. Then the non-null eigenvalues of \mathbf{AB} and \mathbf{BA} are identical.

Lemma 12. Let $\mathbf{A}_1, \dots, \mathbf{A}_k$ be commuting real symmetric matrices of the same order. Then they have a simultaneous spectral decomposition.

4. Non-Negative Definiteness of $\mathbf{Q}^t \mathbf{B} \mathbf{Q}$

Let \mathbf{B} be a real symmetric indefinite matrix of order $n \times n$ and let \mathbf{Q} be a matrix with n rows. In this section, we investigate the conditions under which $\mathbf{Q}^t \mathbf{B} \mathbf{Q}$ is non-negative definite. We shall also give a method of constructing all such matrices \mathbf{Q} .

Let \mathbf{B} be a real symmetric indefinite matrix of order $n \times n$. Let a spectral decomposition of \mathbf{B} be given by

$$\mathbf{B} = (\mathbf{P}_1 : \mathbf{P}_2 : \mathbf{P}_3) \text{diag}(\Delta_1, -\Delta_2, \mathbf{0}) (\mathbf{P}_1 : \mathbf{P}_2 : \mathbf{P}_3)^t,$$

where Δ_i is a positive definite diagonal matrix of order $r_i \times r_i$, $i=1,2$, and $\mathbf{P} = (\mathbf{P}_1 : \mathbf{P}_2 : \mathbf{P}_3)$ is an orthogonal matrix, \mathbf{P}_i being a matrix of $n \times r_i$, $i=1,2,3$ such that $r_1 + r_2 + r_3 = n$.

We prove

Theorem 1. Let \mathbf{B} , \mathbf{P} , Δ_1 , and Δ_2 be as specified above. Let \mathbf{Q} be a matrix of order $n \times s$. Write $\mathbf{R}_i = \mathbf{Q}^t \mathbf{P}_i$, $i = 1,2,3$ and $\mathbf{R} = (\mathbf{R}_1 : \mathbf{R}_2 : \mathbf{R}_3)$. Let $\rho(\mathbf{R}_i) = w_i$, $i = 1,2$. Then $\mathbf{Q}^t \mathbf{B} \mathbf{Q}$ is nnd if and only if there exists a matrix \mathbf{L} with number of columns equal to w_1 such that

$$\mathbf{L} \mathbf{L}^t = \mathbf{R}_1 \Delta_1 \mathbf{R}_1^t$$

and

$$\mathbf{L} \mathbf{A} \mathbf{L}^t = \mathbf{R}_2 \Delta_2 \mathbf{R}_2^t,$$

where \mathbf{A} is a diagonal nnd matrix with exactly w_2 diagonal elements in $(0,1]$ and the rest are equal to 0.

Proof. Notice that $\mathbf{Q} = \mathbf{P} \mathbf{R}^t$ and $\mathbf{Q}^t \mathbf{B} \mathbf{Q} = \mathbf{R} \mathbf{P}^t \mathbf{P} \text{diag}(\Delta_1 : -\Delta_2 : \mathbf{0}) \mathbf{P}^t \mathbf{P} \mathbf{R}^t = \mathbf{R}_1 \Delta_1 \mathbf{R}_1^t - \mathbf{R}_2 \Delta_2 \mathbf{R}_2^t$.

'If part':

$\mathbf{Q}^t \mathbf{B} \mathbf{Q} = \mathbf{R}_1 \Delta_1 \mathbf{R}_1^t - \mathbf{R}_2 \Delta_2 \mathbf{R}_2^t = \mathbf{L} \mathbf{L}^t - \mathbf{L} \mathbf{A} \mathbf{L}^t = \mathbf{L}(\mathbf{I} - \mathbf{A}) \mathbf{L}^t$ is nnd since $\mathbf{I} - \mathbf{A}$ is nnd.

'Only if' part:

Notice that $\mathbf{R}_1\Delta_1\mathbf{R}_1^t$ and $\mathbf{R}_2\Delta_2\mathbf{R}_2^t$ are both nnd. Since $\mathbf{Q}^t\mathbf{B}\mathbf{Q} = \mathbf{R}_1\Delta_1\mathbf{R}_1^t - \mathbf{R}_2\Delta_2\mathbf{R}_2^t$ is nnd, we have $\mathcal{C}(\mathbf{R}_2) \subseteq \mathcal{C}(\mathbf{R}_1)$. By Lemma 5, there exists a nonsingular matrix \mathbf{T} such that

$$\mathbf{R}_1\Delta_1\mathbf{R}_1^t = \mathbf{T} \begin{pmatrix} \Gamma & 0 \\ 0 & 0 \end{pmatrix} \mathbf{T}^t$$

and

$$\mathbf{R}_2\Delta_2\mathbf{R}_2^t = \mathbf{T} \begin{pmatrix} \Omega_1 & 0 \\ 0 & \Omega_2 \end{pmatrix} \mathbf{T}^t$$

where Γ is a positive definite diagonal matrix of order $\omega_1 \times \omega_1$ and $\text{diag}(\Omega_1, \Omega_2)$ is a diagonal nnd matrix of rank ω_2 (Ω_1 is of order $\omega_1 \times \omega_1$).

$$\mathbf{R}_1\Delta_1\mathbf{R}_1^t - \mathbf{R}_2\Delta_2\mathbf{R}_2^t \text{ is nnd} \Leftrightarrow \mathbf{T} \begin{pmatrix} \Gamma & 0 \\ 0 & 0 \end{pmatrix} \mathbf{T}^t - \mathbf{T} \begin{pmatrix} \Omega_1 & 0 \\ 0 & \Omega_2 \end{pmatrix} \mathbf{T}^t \text{ is nnd} \Leftrightarrow \Gamma - \Omega_1 \text{ is nnd and } \Omega_2 = \mathbf{0}.$$

Writing $\mathbf{T} = (\mathbf{T}_1 : \mathbf{T}_2)$ where \mathbf{T}_1 has w_1 columns, we have,

$$\mathbf{R}_1\Delta_1\mathbf{R}_1^t = \mathbf{T}_1\Gamma\mathbf{T}_1^t$$

and

$$\mathbf{R}_2\Delta_2\mathbf{R}_2^t = \mathbf{T}_1\Omega_1\mathbf{T}_1^t$$

Writing $\mathbf{L} = \mathbf{T}_1\Gamma^{\frac{1}{2}}$, we have

$$\mathbf{R}_1\Delta_1\mathbf{R}_1^t = \mathbf{L}\mathbf{L}^t$$

and

$$\mathbf{R}_2\Delta_2\mathbf{R}_2^t = \mathbf{L}\Lambda\mathbf{L}^t,$$

where $\Lambda = \Gamma^{-\frac{1}{2}}\Omega_1(\Gamma^{-\frac{1}{2}})^t$. Further, since $\Gamma - \Omega_1$ is nnd, so is $\mathbf{I} - \Lambda = \Gamma^{-\frac{1}{2}}(\Gamma - \Omega_1)(\Gamma^{-\frac{1}{2}})^t$. Clearly, all diagonal elements of Λ are in $[0, 1]$. Since, $\mathfrak{a}(\Lambda) = \mathfrak{a}(\Omega_1) = \mathfrak{a}(\mathbf{R}_2) = \omega_2$, exactly ω_2 diagonal elements of Λ lie in the interval $(0, 1]$. Q.E.D. \square

Given a real symmetric matrix \mathbf{B} , we now give a method of generating all matrices \mathbf{Q} such that $\mathbf{Q}^t\mathbf{B}\mathbf{Q}$ is nnd. Let \mathbf{B} be as specified just before Theorem 1. Clearly, $\omega_2 \leq \omega_1$. Also, $\omega_2 \leq \rho(\Delta_2)$. Thus, $\omega_2 \leq l = \min\{\omega_1, \rho(\Delta_2)\}$.

We now prove that Algorithm 1 yields the class of all \mathbf{Q} such that $\mathbf{Q}^t\mathbf{B}\mathbf{Q}$ is nnd. First, notice that for each \mathbf{Q} obtained through Algorithm 1,

$$\mathbf{R}_1\Delta_1\mathbf{R}_1^t = \mathbf{L}\mathbf{L}^t$$

and

$$\begin{aligned} \mathbf{R}_2\Delta_2\mathbf{R}_2^t &= \mathbf{LVS}_1 \begin{pmatrix} \mathbf{D}_1^{\frac{1}{2}} & 0 \\ 0 & 0 \end{pmatrix} \mathbf{U}\Delta_2^{-\frac{1}{2}}\Delta_2\Delta_2^{-\frac{1}{2}}\mathbf{U}^t \begin{pmatrix} \mathbf{D}_1^{\frac{1}{2}} & 0 \\ 0 & 0 \end{pmatrix} \mathbf{S}_1^t\mathbf{V}^t\mathbf{L}^t \\ &= \mathbf{LVS}_1 \begin{pmatrix} \mathbf{D}_1 & 0 \\ 0 & 0 \end{pmatrix} \mathbf{S}_1^t\mathbf{V}_1^t\mathbf{L}^t \\ &= \mathbf{L}\Lambda\mathbf{L}^t, \end{aligned}$$

where Λ is $\mathbf{0}$ or is orthogonally similar to a diagonal nnd matrix with at most $l = \min\{\omega_1, \rho(\Delta_2)\}$ diagonal elements in $(0, 1]$. Hence, by Theorem 1, $\mathbf{Q}^t\mathbf{B}\mathbf{Q}$ is nnd.

Next, let $\mathbf{Q}^t\mathbf{B}\mathbf{Q}$ be nnd. Then, there exists $\mathbf{R}_1, \mathbf{R}_2$ such that $\mathbf{Q}^t\mathbf{B}\mathbf{Q} = \mathbf{R}_1\Delta_1\mathbf{R}_1^t - \mathbf{R}_2\Delta_2\mathbf{R}_2^t$ and \mathbf{M} with rank $\rho(\mathbf{R}_1)$ and Λ with rank $\rho(\mathbf{R}_2)$ and Λ such that $\mathbf{M}\mathbf{M}^t = \mathbf{R}_1\Delta_1\mathbf{R}_1^t$ and $\mathbf{M}\Lambda\mathbf{M}^t = \mathbf{R}_2\Delta_2\mathbf{R}_2^t$ where Λ and $\mathbf{I} - \Lambda$ are nnd.

Therefore, by Lemma 2, $\mathbf{R}_1 \Delta_1^{\frac{1}{2}} = \mathbf{M} \mathbf{V}^t$ for some orthogonal matrix \mathbf{V} . Without loss of generality, define $\mathbf{L} = \mathbf{M} \mathbf{V}^t$.

$$\mathbf{L} \mathbf{V} \mathbf{V}^t \mathbf{L}^t = \mathbf{R}_2 \Delta_2 \mathbf{R}_2^t$$

or

$$\mathbf{L} \mathbf{V} \mathbf{B} \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{S}^t \mathbf{V}^t \mathbf{L}^t = \mathbf{R}_2 \Delta_2 \mathbf{R}_2^t$$

where \mathbf{S} is a permutation matrix. So,

$$\mathbf{R}_2 \Delta_2 \mathbf{R}_2^t = \mathbf{L} \mathbf{V} \mathbf{S}_1 \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{S}_1^t \mathbf{V}^t \mathbf{L}^t,$$

where \mathbf{S}_1 is a semi-permutation matrix with $\rho(\Delta_2)$ columns. By Lemma 2,

$$\mathbf{R}_2 \Delta_2^{\frac{1}{2}} = \mathbf{L} \mathbf{V} \mathbf{S}_1 \begin{pmatrix} \mathbf{D}^{\frac{1}{2}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{U}$$

where \mathbf{U} is an orthogonal matrix. Q.E.D.

Algorithm 1 demonstrates how to construct the class of all \mathbf{Q} such that $\mathbf{Q}^t \mathbf{B} \mathbf{Q}$ is nnd in an organized manner. However, it is clear that even when \mathbf{R}_1 is fixed, the class of all \mathbf{Q} such that $\mathbf{Q}^t \mathbf{B} \mathbf{Q}$ is nnd is neither a subspace nor a convex set.

Algorithm 1: The Pochiraju algorithm.

Step 1: Choose \mathbf{R}_1 and \mathbf{R}_3 arbitrarily. (Once \mathbf{R}_1 and \mathbf{R}_3 are chosen and fixed, their ranks ω_1 and ω_3 automatically get fixed.)

Step 2: Construct $\mathbf{L} = \mathbf{R}_1 \Delta_1^{\frac{1}{2}}$.

Step 3: Choose ω_2 arbitrarily such that $0 \leq \omega_2 \leq l = \min\{\omega_1, \rho(\Delta_2)\}$.

Step 4: Choose $\mathbf{D} = \text{diag}(d_1, d_2, \dots, d_{\omega_2}, 0, \dots, 0) = \text{diag}(\mathbf{D}_1 : 0)$ where d_i is an arbitrary number in $(0, 1]$.

Step 5: Construct $\mathbf{R}_2 = \mathbf{R}_1 \Delta_1^{\frac{1}{2}} \mathbf{T} \Delta_2^{-\frac{1}{2}}$ where \mathbf{T} is an arbitrary matrix of rank ω_2 with singular values in $[0, 1]$. (This is actually achieved as follows: Choose \mathbf{S}_1 to be an arbitrary semi-permutation matrix of order $\omega_1 \times \rho(\Delta_2)$, \mathbf{U} be an arbitrary orthogonal matrix of order $\rho(\Delta_2) \times \rho(\Delta_2)$ and construct $\mathbf{R}_2 = \mathbf{L} \mathbf{V} \mathbf{S}_1 \begin{pmatrix} \mathbf{D}_1^{\frac{1}{2}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{U} \Delta_2^{-\frac{1}{2}}$, where \mathbf{V} is an orthogonal matrix.)

Step 6: Construct $\mathbf{Q} = \mathbf{P} \mathbf{R}^t$ where $\mathbf{R} = (\mathbf{R}_1 : \mathbf{R}_2 : \mathbf{R}_3)$.

We now consider two special cases where the construction of the class of all \mathbf{Q} such that $\mathbf{Q}^t \mathbf{B} \mathbf{Q}$ is nnd becomes simple: (i) \mathbf{B} has just one negative eigenvalue and (ii) \mathbf{B} has just one positive eigenvalue.

Case (i): \mathbf{B} has just one negative eigenvalue.

Choose \mathbf{R}_1 and \mathbf{R}_3 arbitrarily. Let $\mathbf{L} \mathbf{L}^t = \mathbf{R}_1 \Delta_1 \mathbf{R}_1^t$ where the number of columns in \mathbf{L} is $\rho(\mathbf{R}_1)$. Since there is only one negative eigenvalue, let us denote it by $-\delta_2$ and the corresponding matrix \mathbf{R}_2 by the vector \mathbf{r}_2 . Now $\mathbf{L} \mathbf{A} \mathbf{L}^t = \delta_2 \mathbf{r}_2 \mathbf{r}_2^t$. Clearly, \mathbf{A} has exactly one nonzero (positive) diagonal element (say λ) which can appear in any of the diagonal elements. So $\mathbf{L} \mathbf{A} \mathbf{L}^t = \lambda \mathbf{l}_i \mathbf{l}_i^t$ where $0 \leq \lambda \leq 1$. So $\mathbf{r}_2 = \sqrt{\frac{\lambda}{\delta_2}} \mathbf{l}_i$. The class of all \mathbf{r}_2 is obtained by choosing λ arbitrarily such that $0 \leq \lambda \leq 1$ and an

arbitrary column (say i th column) \mathbf{l}_i of \mathbf{L} and constructing $\mathbf{r}_2 = \sqrt{\frac{\lambda}{\delta_2}} \mathbf{l}_i$. Then $\mathbf{Q} = \mathbf{P} \begin{pmatrix} \mathbf{R}_1^t \\ \mathbf{r}_2^t \\ \mathbf{R}_3^t \end{pmatrix}$.

As a special instance, we consider the estimated intraclass correlation matrix \mathbf{B} where the estimated intraclass correlation coefficient $\hat{\rho} < \frac{-1}{n-1}$. We now obtain the class of all \mathbf{Q} such that $\mathbf{Q}^t \mathbf{B} \mathbf{Q}$ is nnd.

Here $\Delta_1 = (1 - \hat{\rho}) \mathbf{I}$, $\Delta_2 = -(1 + (n - 1)\hat{\rho})$, $\mathbf{P}_1 \mathbf{P}_1^t = \mathbf{I} - \frac{11^t}{n}$ and $\mathbf{P}_2 \mathbf{P}_2^t = \frac{11^t}{n}$ and \mathbf{P}_3 does not exist since there is no zero eigenvalue.

Choose \mathbf{R}_1 arbitrarily and let \mathbf{L} be a matrix of maximum column rank such that $\mathbf{L} \mathbf{L}^t = (1 - \hat{\rho}) \mathbf{R}_1 \mathbf{R}_1^t$. Choose a column (arbitrarily), say, \mathbf{l}_i and a number λ in the interval $(0, 1]$ (arbitrarily). Construct $\mathbf{r}_2 = \sqrt{\frac{\lambda}{-(1 + (n-1)\hat{\rho})}} \mathbf{l}_i$. Construct $\mathbf{Q} = (\mathbf{P}_1 : \mathbf{P}_2) \begin{pmatrix} \mathbf{R}_1^t \\ \mathbf{r}_2^t \end{pmatrix}$. These are all the matrices \mathbf{Q} such that $\mathbf{Q}^t \mathbf{B} \mathbf{Q}$ is nnd.

Case (ii): \mathbf{B} has just one positive eigenvalue.

Let δ_1 be the positive eigenvalue. Since \mathbf{R}_1 has just one column, we denote it by \mathbf{r}_1 . Choose \mathbf{r}_1 and \mathbf{R}_3 arbitrarily. Denote \mathbf{L} by \mathbf{l} since \mathbf{L} has only one column. Then $\mathbf{l} \mathbf{l}^t = \delta_1 \mathbf{r}_1 \mathbf{r}_1^t$. So, $\mathbf{l} = \sqrt{\delta_1} \mathbf{r}_1$.

Since Λ is a 1×1 matrix, we denote it by λ . As per Theorem 1, $0 \leq \lambda \leq 1$. Choose and fix λ such that $0 \leq \lambda \leq 1$. Then by Theorem 1

$$\lambda \mathbf{l} \mathbf{l}^t = \mathbf{R}_2 \Delta_2 \mathbf{R}_2^t.$$

Notice that

$$\rho(\mathbf{R}_2) = \rho(\mathbf{R}_2 \Delta_2 \mathbf{R}_2^t) = \rho(\lambda \mathbf{l} \mathbf{l}^t) = \rho(\mathbf{l}) = 1.$$

Write $\mathbf{R}_2 = \mathbf{u} \mathbf{v}^t$, where \mathbf{u} and \mathbf{v} are column vectors.

$$\lambda \mathbf{l} \mathbf{l}^t = \mathbf{v}^t \Delta_2 \mathbf{v} \mathbf{u} \mathbf{u}^t.$$

Clearly, \mathbf{u} is a scalar multiple of \mathbf{l} . Choose \mathbf{v} arbitrarily, and then $\mathbf{u} = \sqrt{\frac{\lambda}{\mathbf{v}^t \Delta_2 \mathbf{v}}} \mathbf{l}$. Construct $\mathbf{R}_2 = \mathbf{u} \mathbf{v}^t$ and $\mathbf{Q} = \mathbf{P} \begin{pmatrix} \mathbf{r}_1^t \\ \mathbf{R}_2^t \\ \mathbf{R}_3^t \end{pmatrix}$.

It may be noted that even in these two simple cases, the class of all \mathbf{Q} such that $\mathbf{Q}^t \mathbf{B} \mathbf{Q}$ is nnd is a complex structure. (Neither of them is an affine space).

5. Remarks

In Theorem 1, We have obtained a solution to the following problem in matrix partial orders: Suppose two real symmetric matrices \mathbf{C} and \mathbf{D} are not related by Lowner order. What is the class of all matrices \mathbf{Q} such that $\mathbf{Q}^t \mathbf{C} \mathbf{Q}$ is below $\mathbf{Q}^t \mathbf{D} \mathbf{Q}$ under the Lowner order?

Comparison of the estimators of vector valued parameters is quite common in sample surveys where no estimator is uniformly superior to the others as the difference in the estimated dispersion matrices of the estimators, say, Δ is indefinite (for details, see Section 6.1 of Reference [9]). The results this Section help in identifying the subsets of linear functions of such parameters for which one estimator is superior to the other by finding the class of all \mathbf{Q} for which $\mathbf{Q}^t \Delta \mathbf{Q}$ is nnd.

6. Hausman Test

The usual Hausman Test—in order to test for endogeneity in the random effects model [5], cannot be performed if the difference between the estimated dispersion matrices of the regression coefficient estimators in the fixed effects and random effects models (with homoscedastic structures for the error in the fixed effects model and for the random error and also for the random effects in the random effects model) denoted by Π is not non-negative definite. In this section, we study the difference matrix Π in detail. Since we do not know the regressor matrix at the design stage, we first study when Π is nnd for every choice of the regressor matrix. It turns out that Π is nnd for all regressor matrices \mathbf{X} if and only if the estimated variance of the error in the fixed effects model is at least as

big as the estimator of the variance component of the random noise part in the error in the random effects model. When the difference in the estimated dispersion matrices of the errors in the fixed effects and the random effects models is not nnd, using Algorithm 1, we obtain explicitly the class of all regressor matrices \mathbf{X} for which Π is nnd. Owing to this structure, we show that when the number of regressors is larger than the number of individuals, Π cannot be non-negative definite. Finally, for a given regressor matrix \mathbf{X} , if Π is not nnd, we find an explicit expression for the class of all linear functions of regression coefficients for which the Hausman Test can be performed. We note that Reference [6]'s estimator of the variance component of the random noise part satisfies the property that the estimated variance of the error in the fixed effects model is at least as big as the estimator of the variance component of the random noise part in the error in the random effects model. Thus, with this choice of the estimators of the variance components, Hausman test can be performed for all regressor matrices \mathbf{X} . We observe that we can always get estimators of variance components such that the difference in the dispersion matrices of the error structures in the fixed and random effects models is non-negative definite. Finally, we show that for a suitable choice of the variance component estimators, the difference between the estimated dispersion matrices of the regression coefficient estimators in the fixed effects and random effects models is non-negative definite even when there is heteroscedasticity in the random effects or the random error or both.

We introduce briefly the homoscedastic fixed and random effects panel data models. For details please see References [5,7]. Consider a balanced panel data $(y_{it}, \mathbf{x}_{it}^t), t = 1, \dots, T; i = 1, \dots, N$ where y_{it} is the response and \mathbf{x}_{it}^t is a $1 \times k$ vector of regressor values on k regressors for the i^{th} individual at time point t . Denote $\mathbf{Y}_i = (y_{i1}, \dots, y_{iT})^t$, $\mathbf{X}_i = (\mathbf{x}_{i1}, \dots, \mathbf{x}_{iT})^t$, and $\mathbf{X} = (\mathbf{X}_1^t, \dots, \mathbf{X}_N^t)^t$. Let $\mathbf{1}$ denote a column vector of appropriate order where each component is 1. Denote $\mathbf{F} = \text{diag}(\mathbf{1}, \dots, \mathbf{1})$ where $\mathbf{1}$ is of order $T \times 1$.

The fixed effects specification is given by $\mathbf{Y} = \mathbf{F}\boldsymbol{\alpha} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ where $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_N)^t$ is the vector of fixed effects (treated as non-stochastic), $\boldsymbol{\beta} = (\beta_1, \dots, \beta_k)^t$ is the vector of regression coefficients (also treated as nonstochastic), and $\boldsymbol{\epsilon}$ is a random error vector of order $NT \times 1$ with $E(\boldsymbol{\epsilon}) = \mathbf{0}$ and $D(\boldsymbol{\epsilon}) = \sigma_F^2 \mathbf{I}$. (In the fixed effects model, it is assumed that the observational errors are all uncorrelated and have the same variance, denoted by σ_F^2 .)

The random effects specification is given by $\mathbf{Y} = \mathbf{1}_{NT}\mu + \mathbf{F}\boldsymbol{\alpha} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\xi}$ where $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\xi}$ are as specified in the fixed effects model except that $\boldsymbol{\alpha}$ is treated as random with $E(\boldsymbol{\alpha}) = \mathbf{0}$ and $D(\boldsymbol{\alpha}) = \sigma_\alpha^2 \mathbf{I}$, $\text{cov}(\boldsymbol{\alpha}, \boldsymbol{\xi}) = \mathbf{0}$ and $D(\boldsymbol{\xi}) = \sigma_\xi^2 \mathbf{I}$. We shall denote $\sigma_1^2 = \sigma_\xi^2 + T\sigma_\alpha^2$.

If we denote $\boldsymbol{\eta} = \mathbf{F}\boldsymbol{\alpha} + \boldsymbol{\xi}$ in the random effects model, we get $D(\boldsymbol{\eta}) = \boldsymbol{\Omega} = \sigma_\alpha^2(\mathbf{I}_N \otimes T\mathbf{J}_T) + \sigma_\xi^2(\mathbf{I}_N \otimes \mathbf{I}_T) = \sigma_1^2(\mathbf{I}_N \otimes \mathbf{J}_T) + \sigma_\xi^2(\mathbf{I}_N \otimes \mathbf{E}_T)$.

The usual fixed and random effects estimators of $\boldsymbol{\beta}$, denoted by $\widehat{\boldsymbol{\beta}}_F$ and $\widehat{\boldsymbol{\beta}}_R$ are given by

$$\widehat{\boldsymbol{\beta}}_F = (\mathbf{X}^t(\mathbf{I}_N \otimes \mathbf{E}_T)\mathbf{X})^{-1}\mathbf{X}^t(\mathbf{I}_N \otimes \mathbf{E}_T)\mathbf{Y}$$

and

$$\widehat{\boldsymbol{\beta}}_R = (\mathbf{X}^t\mathbf{E}_{1NT}(\mathbf{E}_{1NT}\boldsymbol{\Omega}\mathbf{E}_{1NT})^{-1}\mathbf{E}_{1NT}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{E}_{1NT}(\mathbf{E}_{1NT}\boldsymbol{\Omega}\mathbf{E}_{1NT})^{-1}\mathbf{E}_{1NT}\mathbf{Y}.$$

Also,

$$D(\widehat{\boldsymbol{\beta}}_F) = (\mathbf{X}^t(\mathbf{I}_N \otimes \mathbf{E}_T)\mathbf{X})^{-1}\sigma_F^2$$

and

$$D(\widehat{\boldsymbol{\beta}}_R) = (\mathbf{X}^t\mathbf{E}_{1NT}(\mathbf{E}_{1NT}\boldsymbol{\Omega}\mathbf{E}_{1NT})^{-1}\mathbf{E}_{1NT}\mathbf{X})^{-1}.$$

Let $s_F^2, s_\alpha^2, s_\xi^2, s_1^2 (= s_\xi^2 + Ts_\alpha^2)$ denote the estimators of $\sigma_F^2, \sigma_\alpha^2, \sigma_\xi^2, \sigma_1^2$ respectively and let $\hat{D}(\widehat{\boldsymbol{\beta}}_F)$ and $\hat{D}(\widehat{\boldsymbol{\beta}}_R)$ denote the estimators of $D(\widehat{\boldsymbol{\beta}}_F)$ and $D(\widehat{\boldsymbol{\beta}}_R)$ where $\sigma_F^2, \sigma_\alpha^2, \sigma_\xi^2, \sigma_1^2$ are replaced by their estimators $s_F^2, s_\alpha^2, s_\xi^2, s_1^2$ respectively. ($\boldsymbol{\Omega}$ is replaced by $\hat{\boldsymbol{\Omega}}$ obtained by plugging in the estimators of the variance components).

As a step towards checking when $\hat{D}(\hat{\beta}_F) - \hat{D}(\hat{\beta}_R)$ is nnd for all \mathbf{X} , we obtain the spectral decomposition of $\mathbf{E}_{1_{NT}}(\mathbf{E}_{1_{NT}}\hat{\Omega}\mathbf{E}_{1_{NT}})^-\mathbf{E}_{1_{NT}}$ in the following Lemma.

Lemma 13. *The spectral decomposition of $\mathbf{E}_{1_{NT}}(\mathbf{E}_{1_{NT}}\hat{\Omega}\mathbf{E}_{1_{NT}})^-\mathbf{E}_{1_{NT}}$ is given by $\frac{1}{s_1^2}(\mathbf{E}_N \otimes \bar{\mathbf{J}}_T) + \frac{1}{s_\xi^2}(\mathbf{I}_N \otimes \mathbf{E}_T)$.*

Proof. Let us start with simplifying $(\mathbf{E}_{1_{NT}}\hat{\Omega}\mathbf{E}_{1_{NT}})$.

First,

$$\begin{aligned}\mathbf{E}_{1_{NT}}\hat{\Omega} &= \{(\mathbf{I}_N \otimes \mathbf{I}_T) - (\bar{\mathbf{J}}_N \otimes \bar{\mathbf{J}}_T)\} \left\{ s_1^2(\mathbf{I}_N \otimes \bar{\mathbf{J}}_T) + s_\xi^2(\mathbf{I}_N \otimes \mathbf{E}_T) \right\} \\ &= s_1^2(\mathbf{I}_N \otimes \bar{\mathbf{J}}_T) + s_\xi^2(\mathbf{I}_N \otimes \mathbf{E}_T) - s_1^2(\bar{\mathbf{J}}_N \otimes \bar{\mathbf{J}}_T), \\ &(\text{since, } (\bar{\mathbf{J}}_N \otimes \bar{\mathbf{J}}_T)(\mathbf{I}_N \otimes \mathbf{E}_T) = \bar{\mathbf{J}}_N \otimes \bar{\mathbf{J}}_T \mathbf{E}_T = \mathbf{0}) \\ &= s_1^2(\mathbf{E}_N \otimes \bar{\mathbf{J}}_T) + s_\xi^2(\mathbf{I}_N \otimes \mathbf{E}_T). \text{ Now,}\end{aligned}$$

$$\begin{aligned}\mathbf{E}_{1_{NT}}\hat{\Omega}\mathbf{E}_{1_{NT}} &= \left\{ s_1^2(\mathbf{E}_N \otimes \bar{\mathbf{J}}_T) + s_\xi^2(\mathbf{I}_N \otimes \mathbf{E}_T) \right\} \{(\mathbf{I}_N \otimes \mathbf{I}_T) - (\bar{\mathbf{J}}_N \otimes \bar{\mathbf{J}}_T)\} \\ &= s_1^2(\mathbf{E}_N \otimes \bar{\mathbf{J}}_T) + s_\xi^2(\mathbf{I}_N \otimes \mathbf{E}_T).\end{aligned}\tag{1}$$

□

Notice that, $\mathbf{E}_N \otimes \bar{\mathbf{J}}_T$ and $\mathbf{I}_N \otimes \mathbf{E}_T$ are both orthogonal projectors and their product is $\mathbf{0}$. Hence, (1) is the spectral decomposition of $\mathbf{E}_{1_{NT}}\hat{\Omega}\mathbf{E}_{1_{NT}}$. One generalized inverse (in fact, the Moore-Penrose inverse) of $\mathbf{E}_{1_{NT}}\hat{\Omega}\mathbf{E}_{1_{NT}}$ is $\frac{1}{s_1^2}(\mathbf{E}_N \otimes \bar{\mathbf{J}}_T) + \frac{1}{s_\xi^2}(\mathbf{I}_N \otimes \mathbf{E}_T)$.

Further, since $\hat{\Omega}$ is positive definite with probability 1, $\mathcal{C}\mathbf{E}_{1_{NT}} \subseteq \mathcal{C}\mathbf{E}_{1_{NT}}\hat{\Omega}\mathbf{E}_{1_{NT}}$ with probability 1. Therefore, $\mathbf{E}_{1_{NT}}\{\mathbf{E}_{1_{NT}}\hat{\Omega}\mathbf{E}_{1_{NT}}\}^-\mathbf{E}_{1_{NT}}$ is invariant under the choices of generalized inverses of $\mathbf{E}_{1_{NT}}\hat{\Omega}\mathbf{E}_{1_{NT}}$.

Now,

$$\begin{aligned}\mathbf{E}_{1_{NT}}\{\mathbf{E}_{1_{NT}}\hat{\Omega}\mathbf{E}_{1_{NT}}\}^-\mathbf{E}_{1_{NT}} &= \\ &\{(\mathbf{I}_N \otimes \mathbf{I}_T) - (\bar{\mathbf{J}}_N \otimes \bar{\mathbf{J}}_T)\} \left\{ \frac{1}{s_1^2}(\mathbf{E}_N \otimes \bar{\mathbf{J}}_T) + \frac{1}{s_\xi^2}(\mathbf{I}_N \otimes \mathbf{E}_T) \right\} \{(\mathbf{I}_N \otimes \mathbf{I}_T) - (\bar{\mathbf{J}}_N \otimes \bar{\mathbf{J}}_T)\} = \\ &\left\{ \frac{1}{s_1^2}(\mathbf{E}_N \otimes \bar{\mathbf{J}}_T) + \frac{1}{s_\xi^2}(\mathbf{I}_N \otimes \mathbf{E}_T) \right\} \{(\mathbf{I}_N \otimes \mathbf{I}_T) - (\bar{\mathbf{J}}_N \otimes \bar{\mathbf{J}}_T)\} = \\ &\frac{1}{s_1^2}(\mathbf{E}_N \otimes \bar{\mathbf{J}}_T) + \frac{1}{s_\xi^2}(\mathbf{I}_N \otimes \mathbf{E}_T). \text{ Q.E.D.}\end{aligned}$$

We are now ready to prove

Theorem 2. *The difference in the estimated dispersion matrices, $\mathbf{\Pi} = \hat{D}(\hat{\beta}_F) - \hat{D}(\hat{\beta}_R)$ is nnd for all \mathbf{X} if and only if $s_F^2 \geq s_\xi^2$.*

Proof. $\hat{D}(\hat{\beta}_F) - \hat{D}(\hat{\beta}_R)$ is nnd, if and only if $[\mathbf{X}^t(\mathbf{I}_N \otimes \mathbf{E}_T)\mathbf{X}]^{-1}s_F^2 - [\mathbf{X}^t\mathbf{E}_{1_{NT}}[\mathbf{E}_{1_{NT}}\hat{\Omega}\mathbf{E}_{1_{NT}}]^{-1}\mathbf{E}_{1_{NT}}\mathbf{X}]^{-1}$ is nnd $\Leftrightarrow \mathbf{X}^t\mathbf{E}_{1_{NT}}[\mathbf{E}_{1_{NT}}\hat{\Omega}\mathbf{E}_{1_{NT}}]^{-1}\mathbf{E}_{1_{NT}}\mathbf{X} - \mathbf{X}^t(\mathbf{I}_N \otimes \mathbf{E}_T)\mathbf{X}\frac{1}{s_F^2}$ is nnd.

But, $\mathbf{E}_{1_{NT}}[\mathbf{E}_{1_{NT}}\hat{\Omega}\mathbf{E}_{1_{NT}}]^{-1}\mathbf{E}_{1_{NT}} - (\mathbf{I}_N \otimes \mathbf{E}_T)\frac{1}{s_F^2} = \frac{1}{s_1^2}(\mathbf{E}_N \otimes \bar{\mathbf{J}}_T) + \frac{1}{s_\xi^2}(\mathbf{I}_N \otimes \mathbf{E}_T) - \frac{1}{s_F^2}(\mathbf{I}_N \otimes \mathbf{E}_T) = \frac{1}{s_1^2}(\mathbf{E}_N \otimes \bar{\mathbf{J}}_T) + \left(\frac{1}{s_\xi^2} - \frac{1}{s_F^2}\right)(\mathbf{I}_N \otimes \mathbf{E}_T)$ (In fact, this is the spectral decomposition.) which is nnd if and only if $\frac{1}{s_\xi^2} - \frac{1}{s_F^2} \geq 0$ or $s_F^2 \geq s_\xi^2$. Q.E.D.

If a computed estimator s_ξ^2 is larger than s_F^2 , it is clear from Theorem 2 that $\mathbf{\Pi} = \hat{D}(\hat{\beta}_F) - \hat{D}(\hat{\beta}_R)$ is not nnd at least for some \mathbf{X} . We now determine the class of all \mathbf{X} for which $\mathbf{\Pi}$ is nnd, so that the Hausman test can be performed for the entire β vector. Towards this end, as we already noted in the proof of Theorem 2, the spectral decomposition of $\mathbf{S} = \mathbf{E}_{1_{NT}}[\mathbf{E}_{1_{NT}}\hat{\Omega}\mathbf{E}_{1_{NT}}]^{-1}\mathbf{E}_{1_{NT}} - (\mathbf{I}_N \otimes \mathbf{E}_T)\frac{1}{s_F^2}$ is given by

$\frac{1}{s_1^2}(\mathbf{E}_N \otimes \bar{\mathbf{J}}_T) + \left(\frac{1}{s_\xi^2} - \frac{1}{s_F^2}\right)(\mathbf{I}_N \otimes \mathbf{E}_T)$. From the spectral decomposition of \mathbf{S} it is clear that the distinct eigen-values of \mathbf{S} are $\frac{1}{s_1^2}, \frac{1}{s_\xi^2} - \frac{1}{s_F^2}$ and 0 with algebraic multiplicities $N - 1, NT - N$ and 1 respectively. Now we can use Algorithm 1 to obtain the class of all \mathbf{X} for which Π is nnd. \square

We prove

Theorem 3. Let $(\mathbf{C}_l, \mathbf{C}_l^t)$ be a rank factorization of \mathbf{E}_l where l is a positive integer. Let $s_\xi^2 > s_F^2$. Then the class of all \mathbf{X} for which Π is nnd is given by

$$\mathbf{X} = \frac{1}{\sqrt{T}}(\mathbf{C}_N \otimes \mathbf{1}_T)\mathbf{R}_1^t + (\mathbf{I}_N \otimes \mathbf{C}_T)\mathbf{R}_2^t + \frac{1}{\sqrt{NT}}(\mathbf{1}_N \otimes \mathbf{1}_T)\mathbf{R}_3^t \text{ where}$$

- (i) \mathbf{R}_1 and \mathbf{R}_3 are arbitrary,
- (ii) $\mathbf{R}_2 = \mathbf{R}_1\mathbf{W}$ where \mathbf{W} is an arbitrary matrix, of rank not greater than that of \mathbf{R}_1 , having singular values in the interval $[0, \theta]$, with $\theta = s_1^2\left(\frac{1}{s_F^2} - \frac{1}{s_\xi^2}\right)$.

Proof. Notice that $\frac{1}{\sqrt{T}}(\mathbf{C}_N \otimes \mathbf{1}_T)$, $\mathbf{I}_N \otimes \mathbf{C}_T$ and $\frac{1}{\sqrt{NT}}(\mathbf{1}_N \otimes \mathbf{1}_T)$ are orthonormal bases of the eigen spaces of $(\mathbf{I} - \mathbf{P}_{\mathbf{1}_{NT}})[(\mathbf{I} - \mathbf{P}_{\mathbf{1}_{NT}})\mathbf{\Omega}(\mathbf{I} - \mathbf{P}_{\mathbf{1}_{NT}})]^{-1}(\mathbf{I} - \mathbf{P}_{\mathbf{1}_{NT}}) - (\mathbf{I}_N \otimes \mathbf{E}_T)\frac{1}{s_F^2}$ corresponding to the eigen-values $\frac{1}{s_1^2}, \frac{1}{s_\xi^2} - \frac{1}{s_F^2}$ and 0 respectively.

Since \mathbf{S} has $NT - N$ negative eigen-values, \mathbf{R}_2 in the expression for \mathbf{X} in Theorem 3 is heavily restricted. Thus, for a large class of matrices \mathbf{X} , we cannot perform the Hausman test for all linear parametric functions. We shall now concentrate on this situation (namely, the difference matrix Π is not nnd) and obtain the class of all linear parametric functions for which we can still perform the Hausman test.

Notice that the Hausman Test can be performed on estimable linear functions, $\mathbf{A}\beta$, if and only if $\mathbf{A}\Pi\mathbf{A}^t$ is non-negative definite. Also, $\mathbf{A}\beta$ are estimable if and only if \mathbf{A} is of the form $\mathbf{A} = \mathbf{Z}(\mathbf{I}_N \otimes \mathbf{E}_T)\mathbf{X}$ for some \mathbf{Z} .

First observe that

$$\begin{aligned} \widehat{D}(\widehat{\beta}_F) - \widehat{D}(\widehat{\beta}_R) &= [\mathbf{X}^t(\mathbf{I}_N \otimes \mathbf{E}_T)\mathbf{X}]^{-1}s_F^2 - [\mathbf{X}^t\left(\frac{1}{s_1^2}(\mathbf{E}_N \otimes \bar{\mathbf{J}}_T) + \frac{1}{s_\xi^2}(\mathbf{I}_N \otimes \mathbf{E}_T)\right)\mathbf{X}]^{-1} \\ &= [\mathbf{X}^t(\mathbf{I}_N \otimes \mathbf{E}_T)\mathbf{X}]^{-1}(s_F^2 - s_\xi^2) + (\mathbf{X}^t(\mathbf{I}_N \otimes \mathbf{E}_T)\mathbf{X})^{-1}\mathbf{X}^t(\mathbf{I}_N \otimes \bar{\mathbf{J}}_T) \\ &= \left[\frac{s_1^2}{s_\xi^2}\mathbf{I}_{NT} + (\mathbf{I}_N \otimes \bar{\mathbf{J}}_T)\mathbf{X}(\mathbf{X}^t(\mathbf{I}_N \otimes \mathbf{E}_T)\mathbf{X})^{-1}\mathbf{X}^t(\mathbf{I}_N \otimes \bar{\mathbf{J}}_T)\right]^{-1}(\mathbf{I}_N \otimes \bar{\mathbf{J}}_T)\mathbf{X}(\mathbf{X}^t(\mathbf{I}_N \otimes \mathbf{E}_T)\mathbf{X})^{-1} \quad (\text{applying} \\ &\quad \text{Lemma 10}). \end{aligned}$$

Consider $\mathbf{A}\beta$. The class of all \mathbf{A} such that the Hausman test can be performed for $\mathbf{A}\beta$ is completely determined by the class of all \mathbf{Z} such that $\mathbf{A} = \mathbf{Z}(\mathbf{I}_N \otimes \mathbf{E}_T)\mathbf{X}$. Write $\mathbf{C} = (\mathbf{I}_N \otimes \mathbf{E}_T)\mathbf{X}$ and $\mathbf{D} = (\mathbf{I}_N \otimes \bar{\mathbf{J}}_T)\mathbf{X}$. In these expressions, \mathbf{C} and \mathbf{D} are the time-varying and time-invariant parts of \mathbf{X} . \square

We want to determine the class of all \mathbf{Z} such that $\mathbf{A}\Pi\mathbf{A}^t$ is nnd.

Now,

$$\mathbf{A}\Pi\mathbf{A}^t = \mathbf{Z}\Phi\mathbf{Z}^t \text{ where}$$

$$\Phi = P_C(s_F^2 - s_\xi^2) + \mathbf{C}(\mathbf{C}^t\mathbf{C})^{-1}\mathbf{D}^t\left(\frac{s_1^2}{s_\xi^2}\mathbf{I} + \mathbf{D}(\mathbf{C}^t\mathbf{C})^{-1}\mathbf{D}^t\right)^{-1}\mathbf{D}(\mathbf{C}^t\mathbf{C})^{-1}\mathbf{C}^t.$$

As before, we need to get the spectral decomposition of Φ in order to determine the class of all \mathbf{Z} such that $\mathbf{Z}\Phi\mathbf{Z}^t$ is nnd. We first prove

Lemma 14. (a) Let $s = \rho(\mathbf{D}(\mathbf{C}^t\mathbf{C})^{-1}\mathbf{D}^t)$. Then $s \leq \rho(\mathbf{C})$.

(b) The non-null eigenvalues of Φ are $(s_F^2 - s_\xi^2) + \left(\frac{s_1^2}{s_\xi^2} + \lambda_i\right)^{-1}\lambda_i, i = 1, \dots, s$ and $s_F^2 - s_\xi^2, i = s + 1, \dots, k$.

Proof. (a) is trivial.

(b) follows from the following facts: (i) Non-null eigenvalues of $\mathbf{B}\mathbf{F}$ and $\mathbf{F}\mathbf{B}$ are the same including multiplicities; (ii) \mathbf{B} and $\mathbf{I} + \mathbf{B}$ commute. (iii) Eigen-values of $\mathbf{f}\mathbf{f}\mathbf{I} + \mathbf{B}$ are obtained by adding α to the eigen-values of \mathbf{B} . (iv) If two real symmetric matrices commute they have simultaneous spectral

decomposition. If \mathbf{B} and \mathbf{F} are two real symmetric matrices of the same order such that $\mathcal{C}(\mathbf{B}) \subseteq \mathcal{C}(\mathbf{F})$, the null space of \mathbf{F} is contained in that of \mathbf{B} . Q.E.D. \square

As a consequence we have the following

Theorem 4. The spectral decomposition of Φ is given by $\Phi = \mathbf{G}\Delta\mathbf{G}^t$ where the columns of \mathbf{G} form an orthonormal basis of \mathcal{C} and Δ is a diagonal matrix whose diagonal entries are the non-null eigen-values of Φ as detailed in Lemma 14.

Using the spectral decomposition of Φ we can determine the class of all \mathbf{Z} such that $\mathbf{Z}\Phi\mathbf{Z}^t$ is nnd using Algorithm 1. From there we can get the class of all $\mathbf{A} = \mathbf{Z}(\mathbf{I}_N \otimes \mathbf{E}_T)\mathbf{X}$ such that we can perform the Hausman test for $\mathbf{A}\beta$.

We now show that there is at least one good estimator of σ_ξ^2 which satisfies Theorem 2.

Notice that $s_F^2 = \frac{R_o^2}{NT-N-k}$ where R_o^2 is the sum of squared residuals in the fixed effects model. Amemiya's estimator of σ_ξ^2 , namely, s_ξ^2 is $\frac{R_o^2}{NT-k-1}$ (see page 16 of Reference [7]). Clearly $s_F^2 \geq s_\xi^2$ for this choice. Amemiya (1971) obtains some optimal properties of this estimator. Thus we proved

Theorem 5. For the choice of the estimators $s_F^2 = \frac{R_o^2}{NT-N-k}$ and $s_\xi^2 = \frac{R_o^2}{NT-k-1}$ of the error variance in the fixed effects model and the variance of the random component of the random effects model respectively, where R_o^2 is the sum of squared residuals in the fixed effects model, the difference in the estimated dispersion matrices of the regression coefficient estimators in the fixed effects and random effects models is non-negative definite.

So far, we considered the case where both the random effects and the random error are homoscedastic. We now examine the case where one or both of them are heteroscedastic. Specifically, we explore whether we can find estimators of the variance components whereby the difference in the dispersion matrices of the design parameter estimators corresponding to the fixed effects and random effects specifications respectively is non-negative definite for all \mathbf{X} .

Let us first write down the fixed effects and random effects specifications with heteroscedasticity.

The fixed effects specification is given by $\mathbf{Y} = \mathbf{F}\alpha + \mathbf{X}\beta + \epsilon$ where $\alpha = (\alpha_1, \dots, \alpha_N)^t$ is the vector of fixed effects (treated as non-stochastic), $\beta = (\beta_1, \dots, \beta_k)^t$ is the vector of regression coefficients (also treated as nonstochastic), and ϵ is a random error vector of order $NT \times 1$ with $E(\epsilon) = 0$ and $D(\epsilon) = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_N^2) \otimes \mathbf{I}_T$.

The random effects specification is given by $\mathbf{Y} = \mathbf{1}_{NT}\mu + \mathbf{F}\alpha + \mathbf{X}\beta + \xi$ where α, β, ξ are as specified in the fixed effects model except that α is treated as random with $E(\alpha) = 0$ and $D(\alpha) = \text{diag}(w_1^2, \dots, w_N^2) \otimes \mathbf{I}_T$, $\text{cov}(\alpha, \xi) = 0$ and $D(\xi) = \text{diag}(r_1^2, \dots, r_N^2) \otimes \mathbf{I}_T$.

If we denote $\eta = \mathbf{F}\alpha + \xi$ in the random effects model, we get $D(\eta) = \Omega = (TD(\alpha) \otimes \mathbf{J}_T) + (D(\xi) \otimes \mathbf{I}_T)$.

Consider the random effects model. Let the estimated dispersion matrices of the random effects and the random error be denoted by $\mathbf{D}_\alpha = \text{diag}(\hat{w}_1^2, \dots, \hat{w}_N^2)$ and $\mathbf{D}_\xi \otimes \mathbf{I}_T$ where $\mathbf{D}_\xi = \text{diag}(\hat{r}_1^2, \dots, \hat{r}_N^2)$.

Hence the estimated error dispersion matrix is $\hat{\Omega} = (TD_\alpha \otimes \mathbf{J}_T) + (\mathbf{D}_\xi \otimes \mathbf{I}_T)$.

Further, the estimated dispersion matrix of the error in the fixed effects specification is $\hat{D}(\epsilon) = \mathbf{D}_F \otimes \mathbf{I}_T$ where $\mathbf{D}_F = \text{diag}(\hat{\sigma}_1^2, \hat{\sigma}_2^2, \dots, \hat{\sigma}_N^2)$.

The fixed and random effects estimators of β , denoted by $\hat{\beta}_F$ and $\hat{\beta}_R$ are given by

$$\hat{\beta}_F = (\mathbf{X}^t(\mathbf{D}_F^{-1} \otimes \mathbf{E}_T)\mathbf{X})^{-1}\mathbf{X}^t(\mathbf{D}_F^{-1} \otimes \mathbf{E}_T)\mathbf{Y}$$

and

$$\hat{\beta}_R = (\mathbf{X}^t\mathbf{E}_{1NT}(\mathbf{E}_{1NT}\hat{\Omega}\mathbf{E}_{1NT})^{-1}\mathbf{E}_{1NT}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{E}_{1NT}(\mathbf{E}_{1NT}\hat{\Omega}\mathbf{E}_{1NT})^{-1}\mathbf{E}_{1NT}\mathbf{Y}.$$

Also,

$$\hat{D}(\hat{\beta}_F) = (\mathbf{X}^t (\mathbf{D}_F^{-1} \otimes \mathbf{E}_T) \mathbf{X})^{-1}$$

and

$$\hat{D}(\hat{\beta}_R) = (\mathbf{X}^t \mathbf{E}_{1_{NT}} (\mathbf{E}_{1_{NT}} \hat{\Omega} \mathbf{E}_{1_{NT}})^{-1} \mathbf{E}_{1_{NT}} \mathbf{X})^{-1}.$$

We now proceed to evaluate the difference in the estimated dispersion matrices of fixed effects and random effects. As before, the difference in the dispersion matrices of the design parameter estimators corresponding to the fixed effects and random effects specifications respectively is non-negative definite for all \mathbf{X} if and only if $(\mathbf{E}_{1_{NT}} (\mathbf{E}_{1_{NT}} \hat{\Omega} \mathbf{E}_{1_{NT}})^{-1} \mathbf{E}_{1_{NT}}) - (\mathbf{D}_F^{-1} \otimes \mathbf{E}_T)$ is nnd. We start with computing

$$\begin{aligned} \mathbf{E}_{1_{NT}} \hat{\Omega} \mathbf{E}_{1_{NT}} &= (\mathbf{I}_N \otimes \mathbf{I}_T - \bar{\mathbf{J}}_N \otimes \bar{\mathbf{J}}_T) (\mathbf{T} \mathbf{D}_\alpha \otimes \bar{\mathbf{J}}_T + \mathbf{D}_\xi \otimes \mathbf{E}_T + \mathbf{D}_\xi \otimes \bar{\mathbf{J}}_T) (\mathbf{I}_N \otimes \mathbf{I}_T - \bar{\mathbf{J}}_N \otimes \bar{\mathbf{J}}_T) \\ &= (\mathbf{T} \mathbf{D}_\alpha \otimes \bar{\mathbf{J}}_T + \mathbf{D}_\xi \otimes \mathbf{E}_T + \mathbf{D}_\xi \otimes \bar{\mathbf{J}}_T - \mathbf{T} \bar{\mathbf{J}}_N \mathbf{D}_\alpha \otimes \bar{\mathbf{J}}_T - \bar{\mathbf{J}}_N \mathbf{D}_\xi \otimes \bar{\mathbf{J}}_T) (\mathbf{I}_N \otimes \mathbf{I}_T - \bar{\mathbf{J}}_N \otimes \bar{\mathbf{J}}_T) \\ &= \mathbf{T} \mathbf{D}_\alpha \otimes \bar{\mathbf{J}}_T + \mathbf{D}_\xi \otimes \mathbf{E}_T + \mathbf{D}_\xi \otimes \bar{\mathbf{J}}_T - \mathbf{T} \bar{\mathbf{J}}_N \mathbf{D}_\alpha \otimes \bar{\mathbf{J}}_T - \bar{\mathbf{J}}_N \mathbf{D}_\xi \otimes \bar{\mathbf{J}}_T - \mathbf{T} \mathbf{D}_\alpha \bar{\mathbf{J}}_N \otimes \bar{\mathbf{J}}_T - \mathbf{D}_\xi \bar{\mathbf{J}}_T \otimes \bar{\mathbf{J}}_T + \\ &\quad \mathbf{T} \bar{\mathbf{J}}_N \mathbf{D}_\alpha \bar{\mathbf{J}}_N \otimes \bar{\mathbf{J}}_T + \bar{\mathbf{J}}_N \mathbf{D}_\xi \bar{\mathbf{J}}_N \otimes \bar{\mathbf{J}}_T \\ &= \mathbf{D}_\xi \otimes \mathbf{E}_T + (\mathbf{T} \mathbf{D}_\alpha + \mathbf{D}_\xi - \mathbf{T} \bar{\mathbf{J}}_N \mathbf{D}_\alpha - \bar{\mathbf{J}}_N \mathbf{D}_\xi - \mathbf{T} \mathbf{D}_\alpha \bar{\mathbf{J}}_N - \mathbf{D}_\xi \bar{\mathbf{J}}_N + \mathbf{T} \bar{\mathbf{J}}_N \mathbf{D}_\alpha + \bar{\mathbf{J}}_N \mathbf{D}_\xi) \otimes \bar{\mathbf{J}}_T \\ &= \mathbf{D}_\xi \otimes \mathbf{E}_T + (\mathbf{T} (\mathbf{E}_N \mathbf{D}_\alpha \mathbf{E}_N) + \mathbf{E}_N \mathbf{D}_\xi \mathbf{E}_N) \otimes \bar{\mathbf{J}}_T = \mathbf{D}_\xi \otimes \mathbf{E}_T + \mathbf{E}_N (\mathbf{T} \mathbf{D}_\alpha + \mathbf{D}_\xi) \mathbf{E}_N \otimes \bar{\mathbf{J}}_T \end{aligned}$$

Now,

$$\begin{aligned} \mathbf{E}_{1_{NT}} (\mathbf{E}_{1_{NT}} \hat{\Omega} \mathbf{E}_{1_{NT}})^{-1} \mathbf{E}_{1_{NT}} &= \{(\mathbf{I}_N \otimes \mathbf{I}_T) - (\bar{\mathbf{J}}_N \otimes \bar{\mathbf{J}}_T)\} \mathbf{D}_\xi^{-1} \otimes \mathbf{E}_T + (\mathbf{E}_N (\mathbf{T} \mathbf{D}_\alpha + \mathbf{D}_\xi) \mathbf{E}_N)^{-1} \otimes \bar{\mathbf{J}}_T \{(\mathbf{I}_N \otimes \mathbf{I}_T - \bar{\mathbf{J}}_N \otimes \bar{\mathbf{J}}_T)\} \\ &= \mathbf{D}_\xi^{-1} \otimes \mathbf{E}_T + (\mathbf{E}_N (\mathbf{T} \mathbf{D}_\alpha + \mathbf{D}_\xi) \mathbf{E}_N)^{-1} \otimes \bar{\mathbf{J}}_T - \bar{\mathbf{J}}_N (\mathbf{E}_N (\mathbf{T} \mathbf{D}_\alpha + \mathbf{D}_\xi) \mathbf{E}_N)^{-1} \otimes \bar{\mathbf{J}}_T \otimes \bar{\mathbf{J}}_T - (\mathbf{E}_N (\mathbf{T} \mathbf{D}_\alpha + \mathbf{D}_\xi) \mathbf{E}_N)^{-1} \bar{\mathbf{J}}_N \otimes \bar{\mathbf{J}}_T + \bar{\mathbf{J}}_N (\mathbf{E}_N (\mathbf{T} \mathbf{D}_\alpha + \mathbf{D}_\xi) \mathbf{E}_N)^{-1} \bar{\mathbf{J}}_N \otimes \bar{\mathbf{J}}_T = \mathbf{D}_\xi^{-1} \otimes \mathbf{E}_T + (\mathbf{E}_N (\mathbf{T} \mathbf{D}_\alpha + \mathbf{D}_\xi) \mathbf{E}_N)^{-1} \mathbf{E}_N \otimes \bar{\mathbf{J}}_T. \end{aligned}$$

Thus, we proved

Theorem 6. $\hat{D}(\hat{\beta}_F) - \hat{D}(\hat{\beta}_R)$ under heteroscedistic specification is nnd if and only if $(\mathbf{D}_\xi^{-1} - \mathbf{D}_F^{-1}) \otimes \mathbf{E}_T + (\mathbf{E}_N (\mathbf{T} \mathbf{D}_\alpha + \mathbf{D}_\xi) \mathbf{E}_N)^{-1} \mathbf{E}_N \otimes \bar{\mathbf{J}}_T$ is nnd.

We can always get a positive definite estimator of the dispersion matrix of random error, that is \mathbf{D}_ξ . For the difference in the estimated dispersion matrices, $\hat{D}(\hat{\beta}_F) - \hat{D}(\hat{\beta}_R)$ to be nnd, we need both $(\mathbf{D}_\xi^{-1} - \mathbf{D}_F^{-1})$ and $(\mathbf{E}_N (\mathbf{T} \mathbf{D}_\alpha + \mathbf{D}_\xi) \mathbf{E}_N)$ to be nnd. We can use Amemiya type estimator to make the first expression to be nnd. It is easy to see that for the second expression to be nnd, it is sufficient that $(\mathbf{T} \mathbf{D}_\alpha + \mathbf{D}_\xi)$ is nnd. This is indeed nnd. (Adapt equation 2.21 of Reference [7] for each individual.) Hence, we can always find error component estimators such that $\hat{D}(\hat{\beta}_F) - \hat{D}(\hat{\beta}_R)$ under heteroscedistic specification is nnd.

The cases where the random error alone or the random effects alone are heteroscedastic are simple special cases of the case we have discussed above. However, if there is heteroscedasticity in the random error, performing the Hausman test requires that T is large, for otherwise the large sample chi-square test will not be valid.

In the next section, we shall show that the problem of finding the class of all \mathbf{Z} such that $\mathbf{Z} \Phi \mathbf{Z}^t$ is nnd is equivalent to solving a quadratic optimization problem.

7. A Quadratic Optimization Problem

In a previous section we obtained the class of all matrices \mathbf{Q} such that $\mathbf{Q}^t \mathbf{B} \mathbf{Q}$ is nnd where \mathbf{B} is a symmetric indefinite matrix. In this section, for a given symmetric indefinite matrix \mathbf{B} , we establish the connection between the following two problems:

- When is $\mathbf{Q}^t \mathbf{B} \mathbf{Q}$ nnd?
- When does $\mathbf{x}^t \mathbf{B} \mathbf{x}$ have a minimum subject to $\mathbf{A} \mathbf{x} = \mathbf{0}$.

We prove

Theorem 7. Let \mathbf{B} be a real symmetric indefinite matrix. Then $\mathbf{Q}^t \mathbf{B} \mathbf{Q}$ is nnd if and only if $\mathbf{x}^t \mathbf{B} \mathbf{x}$ has a minimum subject to $\mathbf{A} \mathbf{x} = \mathbf{0}$ where $\mathcal{N}(\mathbf{A}) = \mathcal{C}(\mathbf{Q})$.

Proof. We note that the orthogonal projectors into $\mathcal{N}(\mathbf{A})$ and $\mathcal{C}(\mathbf{Q})$ are $(\mathbf{I} - \mathbf{A}^+ \mathbf{A})$ and $\mathbf{Q} \mathbf{Q}^+$ respectively. Hence $\mathcal{N}(\mathbf{A}) = \mathcal{C}(\mathbf{Q})$ if and only if $(\mathbf{I} - \mathbf{A}^+ \mathbf{A}) = \mathbf{Q} \mathbf{Q}^+$

‘If part’:

$\mathbf{x}^t \mathbf{B} \mathbf{x}$ has a minimum subject to $\mathbf{A} \mathbf{x} = \mathbf{0}$

$\implies (\mathbf{I} - \mathbf{A}^+ \mathbf{A}) \mathbf{B} (\mathbf{I} - \mathbf{A}^+ \mathbf{A})$ is nnd

$\implies \mathbf{Q} \mathbf{Q}^+ \mathbf{B} \mathbf{Q} \mathbf{Q}^+$ is nnd

$\implies \mathbf{Q}^t \mathbf{B} \mathbf{Q} = \mathbf{Q}^t \mathbf{Q} \mathbf{Q}^+ \mathbf{B} \mathbf{Q} \mathbf{Q}^+ \mathbf{Q}$ is nnd.

‘Only if part’:

$\mathbf{Q}^t \mathbf{B} \mathbf{Q}$ is nnd

$\implies \mathbf{Q} (\mathbf{Q}^t \mathbf{Q})^+ \mathbf{Q}^t \mathbf{B} \mathbf{Q} (\mathbf{Q}^t \mathbf{Q})^+ \mathbf{Q}^t$ is nnd

$\implies \mathbf{Q} \mathbf{Q}^+ \mathbf{B} \mathbf{Q} \mathbf{Q}^+$ is nnd

$\implies \mathbf{x}^t \mathbf{B} \mathbf{x}$ has a minimum subject to $\mathbf{A} \mathbf{x} = \mathbf{0}$.

Given \mathbf{A} , the class $\mathcal{A}_{\mathbf{A}}$ of all \mathbf{Q} such that $\mathcal{N}(\mathbf{A}) = \mathcal{C}(\mathbf{Q})$ can be obtained as follows. Let $(\mathbf{C}, \mathbf{C}^t)$ be a rank factorization of $(\mathbf{I} - \mathbf{A}^+ \mathbf{A})$.

Let $\mathcal{A}_{\mathbf{A}} = \{\mathbf{Q} : \mathbf{Q} = \mathbf{C} \mathbf{T} \text{ where } \mathbf{T} \text{ is an arbitrary full row-rank matrix}\}$. Now, $\mathbf{Q} \in \mathcal{A}_{\mathbf{A}} \implies \mathcal{C}(\mathbf{Q}) = \mathcal{C}(\mathbf{C})$ (since \mathbf{T} is of full row-rank) $= \mathcal{C}(\mathbf{I} - \mathbf{A}^+ \mathbf{A}) = \mathcal{N}(\mathbf{A})$.

Conversely, let $\mathcal{N}(\mathbf{A}) = \mathcal{C}(\mathbf{Q})$. Then $\mathbf{Q} = \mathbf{C} \mathbf{T}$ for some matrix \mathbf{T} . Also $\rho(\mathbf{Q}) = \text{dimension of } \mathcal{N}(\mathbf{A}) = \text{dimension of } \mathcal{C}(\mathbf{I} - \mathbf{A}^+ \mathbf{A}) = \rho(\mathbf{C})$. Since \mathbf{C} has a left inverse, $\rho(\mathbf{Q}) = \rho(\mathbf{T})$. So, (\mathbf{C}, \mathbf{T}) is a rank factorization of \mathbf{Q} . Hence $\mathbf{Q} \in \mathcal{A}_{\mathbf{A}}$. Q.E.D.

Let \mathbf{Q} be a given matrix. We now obtain the class of all matrices \mathbf{A} such that $\mathcal{N}(\mathbf{A}) = \mathcal{C}(\mathbf{Q})$. Let $(\mathbf{D}, \mathbf{D}^t)$ be a rank factorization of $(\mathbf{I} - \mathbf{Q} \mathbf{Q}^+)$. Then the class $\mathcal{D}_{\mathbf{Q}}$ of all matrices \mathbf{A} such that $\mathcal{N}(\mathbf{A}) = \mathcal{C}(\mathbf{Q})$ is given by

$\mathcal{D}_{\mathbf{Q}} = \{\mathbf{A} : \mathbf{A} = \mathbf{W} \mathbf{D}^t \text{ where } \mathbf{W} \text{ is an arbitrary full column-rank matrix}\}$.

Proof follows along similar lines to the earlier case. \square

8. A Quadratic Optimization Problem with Non-Homogeneous Linear Constraints

In this section, we consider the problem of minimization of a quadratic form $\mathbf{x}^t \mathbf{B} \mathbf{x}$ subject to linear constraints $\mathbf{A} \mathbf{x} = \mathbf{b}$, where \mathbf{B} is an $n \times n$ symmetric matrix, \mathbf{A} an $m \times n$ matrix, and \mathbf{b} an $m \times 1$ vector. The case where \mathbf{B} is a pd matrix is well-known [8]. The case where \mathbf{B} is nnd is described in Reference [10]. In this section, for given matrices \mathbf{A} and \mathbf{B} where \mathbf{B} is symmetric (not necessarily nnd), we study when the minimum exists in the following cases:

- (i) For some non-null vector $\mathbf{b} \in \mathbb{R}^m$.
- (ii) For all non-null vectors $\mathbf{b} \in \mathbb{R}^m$.

We shall notice that for a suitable matrix \mathbf{Q} , $\mathbf{Q}^t \mathbf{B} \mathbf{Q}$ being nnd forms an important condition for the existence of a finite solution to the minimization problem. We shall then proceed to characterize the class of all matrices \mathbf{B} and vectors \mathbf{b} (given a matrix \mathbf{A}), such that $\mathbf{x}^t \mathbf{B} \mathbf{x}$ has a finite minimum subject to $\mathbf{A} \mathbf{x} = \mathbf{b}$.

We prove

Theorem 8. Let \mathbf{B} be a real symmetric matrix of order $n \times n$. Let \mathbf{A} be an $m \times n$ matrix and let $\mathbf{b} \in \mathcal{C}(\mathbf{A})$. Consider the minimization problem:

Minimize $\mathbf{x}^t \mathbf{B} \mathbf{x}$ subject to $\mathbf{A} \mathbf{x} = \mathbf{b}$.

Write $\mathbf{H}_1 = (\mathbf{I} - \mathbf{A}^+ \mathbf{A}) \mathbf{B} (\mathbf{I} - \mathbf{A}^+ \mathbf{A})$ and $\mathbf{H}_2 = (\mathbf{I} - \mathbf{A}^+ \mathbf{A}) \mathbf{B} \mathbf{A}^+ \mathbf{A}$.

(a) The problem has a finite solution for some non-null vector \mathbf{b} if and only if

$$\mathbf{H}_1 \text{ is nnd and} \quad (2)$$

$$\mathbf{H}_1 (\mathbf{H}_1 + \mathbf{H}_2 \mathbf{H}_2^t) \mathbf{H}_2 \neq \mathbf{0}. \quad (3)$$

(b) The problem has a finite solution for every $\mathbf{b} \in \mathcal{C}(\mathbf{A})$ if and only if (2) holds and

$$\mathcal{C}(\mathbf{H}_2) \subseteq \mathcal{C}(\mathbf{H}_1). \quad (4)$$

Let $\mathbf{b} = \mathbf{A}\mathbf{u}$. The minimum value in either case is given by

$$\mathbf{u}^t \mathbf{A}^+ \mathbf{A} \mathbf{B} \mathbf{A}^+ \mathbf{A} \mathbf{u} - \mathbf{u}^t \mathbf{H}_2^t \mathbf{H}_1^+ \mathbf{H}_2 \mathbf{u}, \quad (5)$$

and the minimum value is achieved at all vectors \mathbf{x} of the form

$$\mathbf{H}_1^+ \mathbf{H}_2 \mathbf{u} + \mathbf{S}^t \begin{pmatrix} \mathbf{0} \\ \mathbf{1} \end{pmatrix} \quad (6)$$

where $\mathbf{H}_1 = \mathbf{S}^t \begin{pmatrix} \mathbf{\Gamma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{S}$ is a spectral decomposition of \mathbf{H}_1 , \mathbf{S} being orthogonal, $\mathbf{\Gamma}$ a diagonal positive definite matrix, ζ an arbitrary vector in \mathbb{R}^{n-r} , r being the rank of \mathbf{H}_1 .

Proof. By Lemma 3, the class of all \mathbf{x} satisfying $\mathbf{A}\mathbf{x} = \mathbf{b}$ is given by,

$$\mathbf{x} = \mathbf{A}^+ \mathbf{b} + (\mathbf{I} - \mathbf{A}^+ \mathbf{A}) \zeta \quad (7)$$

where ζ is arbitrary. Invoking (7) into $\mathbf{x}^t \mathbf{B} \mathbf{x}$, we get

$$\mathbf{x}^t \mathbf{B} \mathbf{x} = \mathbf{b}^t (\mathbf{A}^+)^t \mathbf{B} \mathbf{A}^+ \mathbf{b} + 2 \mathbf{b}^t (\mathbf{A}^+)^t \mathbf{B} (\mathbf{I} - \mathbf{A}^+ \mathbf{A}) \zeta + \zeta^t (\mathbf{I} - \mathbf{A}^+ \mathbf{A}) \mathbf{B} (\mathbf{I} - \mathbf{A}^+ \mathbf{A}) \zeta. \quad (8)$$

Thus, constrained minimization of $\mathbf{x}^t \mathbf{B} \mathbf{x}$ subject to $\mathbf{A}\mathbf{x} = \mathbf{b}$ is equivalent to unconstrained minimization of the right hand side of (8) over ζ . Since $\mathbf{b} \in \mathcal{C}(\mathbf{A})$, we can write $\mathbf{b} = \mathbf{A}\mathbf{u}$ for some \mathbf{u} . Now, the theorem follows from Lemmas 6 and 7. Q.E.D. \square

Let us identify (2)–(6) in terms of a singular value decomposition of \mathbf{A} . Let us write $\mathbf{b} = \mathbf{A}\mathbf{u}$. Let $\mathbf{A} = \mathbf{U} \begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V}^t$ be a singular value decomposition of \mathbf{A} .

Write $\mathbf{V}^t \mathbf{B} \mathbf{V} = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{pmatrix}$, where \mathbf{R}_{11} is of the same order as Δ and \mathbf{R}_{22} is a square matrix.

It is easy to see that $\mathbf{H}_1 = \mathbf{V} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{22} \end{pmatrix} \mathbf{V}^t$ and $\mathbf{H}_2 = \mathbf{V} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{R}_{21} & \mathbf{0} \end{pmatrix} \mathbf{V}^t$.

Hence, (2) is equivalent to saying that \mathbf{R}_{22} is nnd, (3) is equivalent to saying that $\mathcal{C}(\mathbf{R}_{21}) \cap \mathcal{C}(\mathbf{R}_{22}) \neq \{\mathbf{0}\}$, (4) is equivalent to the statement that $\mathcal{C}(\mathbf{R}_{21}) \subseteq \mathcal{C}(\mathbf{R}_{22})$, and (5) translates to the expression $\mathbf{u}^t \mathbf{V} \begin{pmatrix} \mathbf{R}_{11} - \mathbf{R}_{12} \mathbf{R}_{22}^+ \mathbf{R}_{21} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V}^t \mathbf{u}$.

Let $\rho(\mathbf{R}_{22}) = s$, and let $\mathbf{R}_{22} = \mathbf{M} \begin{pmatrix} \mathbf{\Gamma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{M}^t$ be a spectral decomposition of \mathbf{R}_{22} , where \mathbf{M} is orthogonal and $\mathbf{\Gamma}$ is a diagonal positive definite matrix of order $s \times s$. Then a spectral decomposition

of \mathbf{H}_1 is given by $\mathbf{V} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{M} \end{pmatrix} \mathbf{P} \begin{pmatrix} \mathbf{\Gamma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{P}^t \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}^t \end{pmatrix} \mathbf{V}^t$, where \mathbf{P} is a suitable permutation matrix.

In view of this, (6) translates to the expression $\mathbf{V} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{R}_{22}^+ \mathbf{R}_{21} & \mathbf{0} \end{pmatrix} \mathbf{V}^t \mathbf{u} + \mathbf{V} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{M} \end{pmatrix} \mathbf{P} \begin{pmatrix} \mathbf{0} \\ \boldsymbol{\zeta} \end{pmatrix}$, where $\boldsymbol{\zeta}$ is an arbitrary vector in \mathbb{R}^{n-s} .

We note that (2), namely \mathbf{H}_1 should be nnd, is a key factor for the existence of a finite solution to the constrained optimization problem under consideration, which falls into the line of investigation in Section 3.

Let \mathbf{A} be a given $m \times n$ matrix of rank r . The above identification helps us in characterizing the class of all real symmetric matrices \mathbf{B} and the class of all vectors $\mathbf{b} \in \mathcal{C}(\mathbf{A})$ such that the constrained optimization problem

Minimize $\mathbf{x}^t \mathbf{B} \mathbf{x}$ subject to $\mathbf{A} \mathbf{x} = \mathbf{b}$

has a finite solution.

As before, let $\mathbf{A} = \mathbf{U} \begin{pmatrix} \boldsymbol{\Delta} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V}^t$ be singular value decomposition of \mathbf{A} , where \mathbf{U} and \mathbf{V} are orthogonal matrices and $\boldsymbol{\Delta}$ is a positive definite diagonal matrix of order $r \times r$. Write $\mathbf{V}^t \mathbf{B} \mathbf{V} = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{pmatrix}$, where \mathbf{R}_{11} is of order $r \times r$ and \mathbf{R}_{22} is of order $(n-r) \times (n-r)$. Since, $\mathbf{b} \in \mathcal{C}(\mathbf{A})$, write $\mathbf{b} = \mathbf{A} \mathbf{u}$. Characterizing \mathbf{B} and \mathbf{b} is equivalent to characterizing \mathbf{R}_{11} , \mathbf{R}_{21} , \mathbf{R}_{22} , and \mathbf{u} .

If the minimization problem should have a finite solution for every $\mathbf{b} \in \mathcal{C}(\mathbf{A})$, then the class of all

\mathbf{B} is given by $\mathbf{B} = \mathbf{V} \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{21}^t \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{pmatrix} \mathbf{V}^t$.

where (a) \mathbf{R}_{11} is an arbitrary real symmetric matrix of order $r \times r$

(b) \mathbf{R}_{22} is an arbitrary nnd matrix of order $(n-r) \times (n-r)$

(c) $\mathbf{R}_{21} = \mathbf{R}_{22} \mathbf{D}$, where \mathbf{D} is an arbitrary matrix of order $(n-r) \times r$.

If the minimization problem should have a solution for some non-null vector $\mathbf{b} \in \mathcal{C}(\mathbf{A})$, then the class of all \mathbf{B} is given by $\mathbf{B} = \mathbf{V} \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{21}^t \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{pmatrix} \mathbf{V}^t$, satisfying (a) and (b) as above and

(d) $((\mathbf{F} \mathbf{y} : \mathbf{G}), \mathbf{W})$ is a rank-factorization of \mathbf{R}_{22} , where

(i) $(\mathbf{F}, \mathbf{F}^t)$ is a rank factorization of \mathbf{R}_{22} ,

(ii) \mathbf{y} is arbitrary non-null vector,

(iii) \mathbf{G} is arbitrary such that $(\mathbf{F} \mathbf{y} : \mathbf{G})$ is of full column rank, and

(iv) \mathbf{W} is an arbitrary full row rank matrix.

The class of all \mathbf{b} is obtained as follows:

Using \mathbf{B} as obtained above, compute $\mathbf{J} = \mathbf{H}_1(\mathbf{H}_1 + \mathbf{H}_2 \mathbf{H}_2^t)^- \mathbf{H}_2 \mathbf{H}_2^t$. Let \mathbf{w} be an arbitrary non-zero vector in $\mathcal{C}(\mathbf{J})$. Let \mathbf{u} be an arbitrary solution of $\mathbf{H}_2 \mathbf{u} = \mathbf{w}$. Compute $\mathbf{b} = \mathbf{A} \mathbf{u}$. Notice that $\mathbf{b} \neq \mathbf{0}$, for, if $\mathbf{b} = \mathbf{0}$, $\mathbf{A} \mathbf{u} = \mathbf{0}$, and hence $\mathbf{H}_2 \mathbf{u} = \mathbf{0}$, which is a contradiction, since $\mathbf{H}_2 \mathbf{u} = \mathbf{w} \neq \mathbf{0}$.

We prove

Theorem 9. Let \mathbf{B} be a symmetric indefinite matrix of order $n \times n$, and \mathbf{A} be an $n \times n$ matrix, and \mathbf{b} be an $m \times 1$ vector in the column space of \mathbf{A} . If $\mathbf{x}^t \mathbf{B} \mathbf{x}$ has a finite minimum subject to $\mathbf{A} \mathbf{x} = \mathbf{b}$ for every $\mathbf{b} \in \mathcal{C}(\mathbf{A})$, then there exists a generalized inverse \mathbf{G} of \mathbf{A} such that

$$(\mathbf{I} - \mathbf{G} \mathbf{A}) \mathbf{B} (\mathbf{I} - \mathbf{G} \mathbf{A}) \text{ is nnd} \quad (9)$$

$$\text{and } \mathbf{B} \mathbf{G} \mathbf{A} = (\mathbf{G} \mathbf{A})^t \mathbf{B}. \quad (10)$$

Proof. Since $\mathbf{x}^t \mathbf{B} \mathbf{x}$ has a finite minimum subject to $\mathbf{A} \mathbf{x} = \mathbf{b}$ for every \mathbf{b} , by Lemma 13, we have, $(\mathbf{I} - \mathbf{A}^+ \mathbf{A}) \mathbf{B} (\mathbf{I} - \mathbf{A}^+ \mathbf{A})$ is nnd. Whenever, \mathbf{G} is a generalized inverse of \mathbf{A} , $(\mathbf{I} - \mathbf{A}^+ \mathbf{A})(\mathbf{I} - \mathbf{G} \mathbf{A}) =$

$\mathbf{I} - \mathbf{GA}$. Now, it follows that, $(\mathbf{I} - \mathbf{GA})^t(\mathbf{I} - \mathbf{A}^+\mathbf{A})\mathbf{B}(\mathbf{I} - \mathbf{A}^+\mathbf{A})(\mathbf{I} - \mathbf{GA}) = (\mathbf{I} - \mathbf{GA})^t\mathbf{B}(\mathbf{I} - \mathbf{GA})$ is nnd. Further, from the discussion after Theorem 7, it follows that \mathbf{R}_{22} is nnd and $\mathcal{C}(\mathbf{R}_{21}) \subseteq \mathcal{C}(\mathbf{R}_{22})$.

Since \mathbf{R}_{22} is nnd, $(\mathbf{I} - \mathbf{GA})^t\mathbf{B}(\mathbf{I} - \mathbf{GA}) = \mathbf{V} \begin{pmatrix} -\Delta\mathbf{M}^t \\ -\mathbf{I} \end{pmatrix} \mathbf{R}_{22} \begin{pmatrix} \mathbf{M}\Delta & -\mathbf{I} \end{pmatrix} \mathbf{V}^t$ is nnd, whatever \mathbf{M} be. Further, since $\mathcal{C}(\mathbf{R}_{21}) \subset \mathcal{C}(\mathbf{R}_{22})$, there exists a matrix \mathbf{T} such that $\mathbf{R}_{21} = \mathbf{R}_{22}\mathbf{T}$. Write $\mathbf{M} = \mathbf{T}\Delta^{-1}$. Now it is easy to verify that, for this choice of \mathbf{M} , $\mathbf{G} = \mathbf{V} \begin{pmatrix} \Delta^{-1} & \mathbf{L} \\ \mathbf{M} & \mathbf{N} \end{pmatrix} \mathbf{U}^t$ is a generalized inverse of \mathbf{A} (where \mathbf{L} and \mathbf{N} are arbitrary) such that $\mathbf{BGA} = (\mathbf{GA})^t\mathbf{B}$. Q.E.D. \square

Remark 1. \mathbf{A}^+ does not in general have the properties (9) and (10).

Is there anything special about a generalized inverse \mathbf{G} of \mathbf{A} satisfying (9) and (10)?

It turns out that every generalized inverse \mathbf{G} of \mathbf{A} satisfying (9) and (10) is in fact a minimum semi-norm generalized inverse of \mathbf{A} under a suitable semi-inner product. To see this construct

$$\mathbf{S} = \mathbf{V} \begin{pmatrix} k\Delta^{-2} - \mathbf{R}_{11} & 0 \\ 0 & 0 \end{pmatrix} \mathbf{V}^t$$

where $k > \text{tr}(\mathbf{R}_{12}^t \mathbf{R}_{22}^- \mathbf{R}_{21})$. Then,

$$\mathbf{B} + \mathbf{A}^t\mathbf{SA} = \mathbf{V} \begin{pmatrix} k\mathbf{I} & \mathbf{R}_{21}^t \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{pmatrix} \mathbf{V}^t$$

is nnd, since \mathbf{R}_{22} is nnd, $\mathcal{C}(\mathbf{R}_{21}) \subset \mathcal{C}(\mathbf{R}_{22})$ and $k\mathbf{I} - \mathbf{R}_{21}^t \mathbf{R}_{22}^- \mathbf{R}_{21}$ is nnd (by virtue of choice of k and Lemma 1). Also, $\mathbf{AGA} = \mathbf{A}$ and

$$(\mathbf{B} + \mathbf{A}^t\mathbf{SA})\mathbf{GA} = \mathbf{BGA} + \mathbf{A}^t\mathbf{SA} = (\mathbf{GA})^t\mathbf{B} + \mathbf{A}^t\mathbf{SA},$$

and is thus symmetric. Hence, \mathbf{G} is a minimum semi-norm generalized inverse of \mathbf{A} under the semi-inner product $(\mathbf{x}, \mathbf{y}) = \mathbf{y}^t(\mathbf{B} + \mathbf{A}^t\mathbf{SA})\mathbf{x}$ (see Theorem 1.4 of Reference [11]).

9. Conclusions, Limitations and Future Research

In this article we discussed extensively and provided new results for the optimization of quadratic structures and respective corrections to the construction of covariance matrices. This is an area has a long history in econometrics and financial economics with various applications in testing panel data models, handling volatility, to applications in portfolio management. Furthermore, the implications in forecasting are numerous as all relate to decision-making at some future time period: panel data models can be used to generate out-of-sample forecasts; a covariance matrix can be used in optimizing a portfolio or in a model for volatility and correlation forecasting. The economic implications of all the aforementioned are profound. We have presented a novel statistical approach for solving a particular class of quadratic optimization problems. At the center of the mathematical derivations is a symmetric indefinite matrix, the indefinite nature of which can come from many sources usually rank deficiency or rank indeterminacy driven from redundant information in the variables from which we compute that said matrix. The problem of this indeterminacy leads to subsequent problems in the context of: a matrix-version of the well-known Hausman test in econometrics; a large portfolio optimization the covariance matrix of the financial returns might not be positive definite. This is the body of literature and application we contribute to.

As any other statistical derivation, this paper comes with the usual limitations and caveats of any statistical analysis—we do provide solutions to well known problems, but for which there are alternative solutions [6], and as such no solution is universally better, and given collected samples the

researcher may have to try an extensive array of available tools, to which we contribute emphatically one more here.

For future research, we leave the investigation of further application areas for our propositions, as well as simulations for a wide range of panel data, and optimization problems.

Author Contributions: Conceptualization, B.P.; methodology, B.P. and S.S.; software, D.D.T.; validation, B.P., S.S. and D.D.T.; formal analysis, B.P. and S.S.; investigation, B.P., S.S., D.D.T.; resources, B.P. and S.S.; data curation, S.S.; writing—original draft preparation, B.P.; writing—review and editing, S.S., D.D.T., and K.N.; supervision, B.P.; project administration, B.P. and K.N.; All authors have read and agreed to the published version of the manuscript. (S.S. and K.N. on behalf of B.P.)

Funding: This research received no external funding.

Acknowledgments: Bhimasankaram Pochiraju has acknowledged and Sridhar Seshadri gratefully acknowledges the research support from the Applied Statistics and Computing Lab, Indian School of Business. Nikolopoulos gratefully acknowledges the research support from Indian School of Business during his two visits in the campuses in Hyderabad in 2017, and the extended stay in Mohali in early 2018 while delivering the PGP elective module on ‘Forecasting Analytics’.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Linton, O.; Tang, H. *Estimation of the Kronecker Covariance Model by Quadratic Form*; Cambridge Working Papers in Economics 2050; Faculty of Economics, University of Cambridge: Cambridge, UK, 2020.
2. Eriksson, A.; Preve, D.; Yu, J. Forecasting Realized Volatility Using a Nonnegative Semiparametric Modell. *J. Risk Financ. Manag.* **2019**, *12*, 139. [CrossRef]
3. Toloo, M.; Mensah, E. Robust optimization with nonnegative decision variables: A DEA approach. *Comput. Ind. Eng.* **2019**, *127*, 313–325. [CrossRef]
4. Mitra, S.; Bhimasankaram, P.; Malik, S. *Matrix Partial Orders. Shorted Operators and Applications*; World Scientific: Singapore, 2010.
5. Greene, W. *Econometric Analysis*, 7th ed.; Prentice Hall: Upper Saddle River, NJ, USA, 2012.
6. Amemiya, T. The estimation of variance components in a variance components model. *Int. Econ. Rev.* **1971**, *12*, 1–13. [CrossRef]
7. Baltagi, B. *Econometric Analysis of Panel Data*, 5th ed.; Wiley: New York, NY, USA, 2013.
8. Rao, A.; Bhimasankaram, P. *Linear Algebra*, 2nd ed.; Hindustan Book Agency: New Delhi, India, 2000.
9. Goga, C. Variance Estimators in Survey Sampling. Available online: http://goga.perso.math.cnrs.fr/ChapVar1_coursBesan.pdf (accessed on 24 March 2008).
10. Kambo, N. *Mathematical Programming Techniques*; Affiliated East West Press: New Delhi, India, 1984.
11. Rao, C.; Mitra, S. *Generalized Inverse of Matrices and Its Applications*; Wiley: New York, NY, USA, 1971.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).